

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
À L'OBTENTION DE LA
MAÎTRISE EN MATHÉMATIQUE ET INFORMATIQUE APPLIQUÉES

PAR
JONATHAN BOISCLAIR

INTÉGRATION DE L'ANALYSE DE SIGNAUX BIOMÉTRIQUES DANS UN ENVIRONNEMENT
DE RÉALITÉ VIRTUELLE POUR LA DÉTECTION PAR APPRENTISSAGE AUTOMATIQUE
DES FACULTÉS D'UNE PERSONNE

DÉCEMBRE 2019

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

SOMMAIRE

Afin d'évaluer les facultés d'une personne, il existe divers tests symptomatiques comme les tests du nystagmus, de convergence oculaire et de réaction à la lumière des pupilles. Ces tests sont effectués manuellement et sont essentiellement appliqués aux yeux. L'objectif de cette recherche est d'améliorer un prototype automatique déjà développé pour automatiser ces tests symptomatiques et aussi intégrer une approche d'apprentissage automatique. Ce nouveau prototype plus compact développé à l'intérieur d'un environnement de réalité virtuelle est aussi constitué d'un système d'analyse d'images vidéo, mais avec des fonctionnalités optimisées. Ce dernier prototype comme les premiers, effectue aussi la capture d'images vidéo du mouvement et des réactions de l'œil soumis à des stimuli visuels. Le traitement de ces séquences d'images permet de faire le suivi du déplacement des pupilles ainsi que leurs réactions aux changements d'illumination ambiante et d'ainsi reconstituer les signaux correspondant au déplacement des yeux et des dimensions des pupilles. L'analyse de ces signaux et d'autres signaux biométriques permet d'identifier, entre autres, les saccades correspondant au nystagmus horizontal, le degré de convergence des yeux ainsi que les dimensions des pupilles sous divers modes d'illumination permettant d'évaluer les facultés d'une personne. Le processus d'évaluation des facultés affaiblies sera basé sur une approche d'apprentissage automatique dans laquelle la connaissance est modélisée sous forme d'arbre qui à partir des différents résultats obtenus des tests symptomatiques pourra déduire le type d'intoxication d'une personne. Ce prototype est testé sur des volontaires ayant des facultés affaiblies par la consommation de diverses substances ou par la fatigue extrême. L'efficacité du nouvel appareil et des tests automatisés pour déterminer l'état des facultés d'un individu est ensuite examinée.

ABSTRACT

To assess a person's faculties, there are various symptomatic tests such as the horizontal gaze nystagmus test, eye convergence and pupil light reaction tests. These tests are performed manually and are essentially applied to the eyes. The objective of this research is to improve an automatic prototype already developed to automate these symptomatic tests and to integrate a self-learning approach. This new, more compact prototype developed within a virtual reality environment, is also made up of a video image analysis system but with optimized features. The latter prototypes, like the first developed, also captures video images of the movement and the reactions of the eye subjected to visual stimulus. The processing of these sequences of images allows the tracking of the displacement of pupils as well as their reactions to changes of ambient illumination and thus the reconstruction of the signals corresponding to the displacement of the eyes and dimensions of pupils. The analysis of these signals and other biometric signals allows the identification, among other things, the saccades corresponding to the horizontal nystagmus, the degree of convergence of the eyes as well as the dimensions of the pupils under various modes of illumination allowing the assessment of a person's faculties. The impaired assessment process will be based on a self-learning approach in which knowledge is modelled as a tree constructed from the different results obtained from the symptomatic tests from which one can infer the type of intoxication of a person. This prototype is tested on volunteers who are impaired by the consumption of various substances or by extreme fatigue. The effectiveness of the new device and automated tests to determine the state of an individual's faculties is then examined.

AVANT PROPOS ET REMERCIEMENTS

Ce projet a pour objectif de mettre en œuvre un sujet qui sera toujours en évolution et qui cherche encore des outils dans le monde réel. Le but ultime est la sécurité du public dans différentes sphères de la société telles que la conduite automobile, l'utilisation de machinerie lourde ou toute autre tâche nécessitant beaucoup d'attention.

La sécurité est quelque chose qui me tient à cœur, car on ne sait jamais ce qui pourrait arriver à nos êtres chers. Pouvoir savoir qu'on va les revoir le lendemain nous enlève une grande charge. On parle ici de la sécurité dans tous les aspects de la vie. Un accident est vite arrivé, alors pouvoir le prévenir ne serait pas un luxe.

Je remercie tous ceux qui m'ont aidé au long de ce court chemin et qui m'ont permis de terminer cette œuvre, principalement ma famille, mon directeur François Meunier et tous les professeurs du département de Mathématiques et d'informatique.

TABLE DES MATIÈRES

SOMMAIRE.....	I
ABSTRACT	II
AVANT PROPOS ET REMERCIEMENTS.....	III
TABLE DES MATIÈRES.....	IV
LISTE DES FIGURES	VII
LISTE DES TABLEAUX	IX
CHAPITRE 1 INTRODUCTION	1
1.1 Objectifs de ce mémoire	2
CHAPITRE 2 REVUE DE LITTÉRATURE.....	4
2.1 Recherches antérieures	4
2.1.1 Premier prototype.....	4
2.1.2 Second prototype.....	5
2.1.3 Quatrième prototype	6
2.1.4 Premier prototype sur plateforme mobile.....	7
2.2 Revues de littérature reliée au sujet de recherche	7
2.2.1 Anatomie de l'œil	7
2.2.2 Tests standards	8
2.2.3 Analyse des yeux par ordinateur.....	12
2.2.4 Alcoolémie	13
2.2.5 Autres drogues.....	15
2.2.6 Traitement d'image avec Emgu.CV	16
2.2.7 Traitement d'image avec ImageProcessing.....	17
2.3 Revue de la littérature reliée directement aux nouvelles spécialités de cette recherche	19
2.3.1 N-Gram.....	19

2.3.2 Représentation des N-Gram en mémoire	21
2.4 Algorithme de représentation des résultats en arbre et comparaison	23
2.4.1 Introduction à l'algorithme en arbre	23
2.4.2 Résumé	23
2.4.3 Problématique	24
2.4.4 Méthodologie (Les étapes)	24
2.4.5 Amélioration de l'algorithme en arbre	34
2.4.6 Différences majeures par rapport à l'algorithme des N-Gram	35
2.4.7 Généralisation de l'algorithme	35
2.5 Conclusion	36
CHAPITRE 3 MÉTHODE EXPÉRIMENTALE	37
3.1 Introduction et hypothèses de base	37
3.2 Application de l'algorithme en arbre à des données de type autres que texte	38
3.3 Méthode de tests des différentes solutions possibles	39
3.3.1 Méthode d'acquisition des données	39
3.3.2 Méthode d'identification des différentes solutions	41
3.4 Calcul et choix de la meilleure solution	44
3.4.1 Différents calculs possibles	44
3.5 Améliorations au dispositif de capture	50
3.6 Conclusion	61
CHAPITRE 4 RÉSULTATS, DISCUSSION ET INTERPRÉTATION	62
4.1 Extraction des données à partir de vidéos	63
4.2 Résultats de l'apprentissage	69
4.3 Discussion des résultats	81
4.4 Avancées possibles	82
4.5 Conclusion	83
CHAPITRE 5 CONCLUSION	84

BIBLIOGRAPHIE.....	86
--------------------	----

LISTE DES FIGURES

Figure 1 Prototype utilisé par Pierre Drolet [4]	4
Figure 2 Prototype utilisé par Pier-Olivier Houde [5]	5
Figure 3 Processus extraction des pupilles introduit par Dominic Adam [6]	6
Figure 4 Anatomie de l'œil humain Récupéré sur Guide Vue [9]	7
Figure 5 Test du Nystagmus horizontal droit.....	9
Figure 6 Images capturées pendant le test du Nystagmus horizontal.	9
Figure 7 Courbe de dilatation/contraction	10
Figure 8 Images capturées lors d'un test de dilatation	11
Figure 9 Images capturées lors d'un test de convergence et déplacement du stimulus .	11
Figure 10 Image d'un œil.....	12
Figure 11 Segmentation en Open CV.....	18
Figure 12 Exemple de N-Gram en TLN	19
Figure 13 Algorithme N-GRAM.....	20
Figure 14 Représentations de N-Gram en plusieurs tableaux.....	22
Figure 15 Représentations de N-Gram en tableau	22
Figure 16 Création d'arbres simple	25
Figure 17 Représentations en arbre d'une phrase	26
Figure 18 Meilleur des cas - Arbre.....	27
Figure 19 Cas moyen - Arbre	27
Figure 20 Pire des cas - Arbre.....	28
Figure 21 Additions d'arbre.....	30
Figure 22 Soustractions d'arbre.....	32
Figure 23 Soustractions d'arbre de taille différente	33
Figure 24 Exemples de mots égaux, similaires et différents	38
Figure 25 Processus d'identification d'un test	42
Figure 26 Casque Oculus.....	51
Figure 27 Socles pour caméras avant impression	52
Figure 28 Dispositif de capture de séquences vidéo par caméra NIR.....	52
Figure 29 Processus d'identification de la pupille dans les vidéos	54
Figure 30 Processus extraction de l'Iris	55
Figure 31 Formule d'identification de l'angle de visée corrigée	56

Figure 32 Capteur de données musculaire (EMG).....	57
Figure 33 SI8902B récupérée de Silicon Labs [27].....	58
Figure 34 Capteur cardiaque [28]	59
Figure 35 Carte pour accueillir capteur de température	59
Figure 36 PCB acquisition	60
Figure 37 Résultats d'un test de convergence de l'œil gauche	63
Figure 38 Résultats d'un test de dilatation de l'œil gauche	65
Figure 39 Résultats d'un test de nystagmus de l'œil gauche	66
Figure 40 Résultats d'un test de nystagmus de l'œil gauche(Suite).....	67
Figure 41 Classification selon CA.....	76
Figure 42 Classification selon CB	77
Figure 43 Classification selon CC.....	78
Figure 44 Classification selon CD.....	79
Figure 45 Classification du nombre de modes par la qualité de la classification en % ...	79
Figure 46 Images d'un œil flou ou couvert.....	81

LISTE DES TABLEAUX

Tableau 1 Choix CA (Point de coupe)	45
Tableau 2 Choix CB (Degré de liberté)	46
Tableau 3 Choix CC (unification des différentes sources de données)	46
Tableau 4 Les différentes données provenant d'un vidéo	48
Tableau 5 Exemple de points selon la position	49
Tableau 6 Choix CD (Unification des tests)	49
Tableau 7 Résultat d'identification	71
Tableau 8 Validations selon les 4 modes (Partiel 83% et plus)	73
Tableau 9 Validation croisée	74
Tableau 10 Matrice de confusion	74

CHAPITRE 1

INTRODUCTION

Dans le domaine de la conduite automobile, beaucoup de facteurs affectent la conduite [1]. On entend souvent parler de facultés affaiblies. Un concept assez inclusif et en même temps vague, alcool, drogues, fatigue, Qui sont des facteurs pouvant affecter nos facultés lors de la conduite automobile. Pour chaque type de réduction des facultés, ils existent des tests (DRE) [2], voire des exercices à réaliser. Ces tests sont des tests standards défini pour identifier les facultés d'un individu. Cependant, dû à la nature différente des humains, ces tests ne sont pas absolus et peuvent donner des résultats erronés. Dans bien des domaines, quand il n'y a pas homogénéisation de la population, un algorithme d'apprentissage automatique permet de mieux identifier la présence des facultés affaiblies que les tests traditionnels [3]. L'apprentissage automatique, un terme assez à la mode, peut inclure beaucoup d'algorithmes différents, par exemple les réseaux de neurones, les réseaux bayésiens, les algorithmes K-Means et X-Means, ayant chacun leurs forces et leurs faiblesses. Certains ont un nombre d'entrées et donne des résultats précis, mais ne peuvent pas s'éloigner de l'ensemble d'apprentissage, d'autres prédisent beaucoup mieux lorsqu'on s'éloigne de l'apprentissage, mais catégorisent moins bien l'ensemble même. Certains s'appliquent au texte, certains aux calculs, d'autres aux images.

Les différents tests actuellement réalisés pour détecter les facultés affaiblies prennent beaucoup de temps, et ne peuvent pas tous être réalisés au moment désiré, soit au moment où un policier intercepte un conducteur. Certains tests, comme les prises de sang et d'urine doivent être envoyées au laboratoire pour analyse ou être faits en milieu contrôlé au poste. Les tests de dilatation fait pendant le jour doivent présentement être effectué au poste dans une pièce sombre. La possibilité d'avoir une machine portable capable de faire certains de ces tests dans un environnement "contrôlé" virtuellement, permettrait d'avoir de meilleurs résultats et cela plus rapidement. En ce qui a trait à cet environnement contrôlé, le meilleur moyen est d'utiliser un environnement totalement virtuel tel qu'un casque de réalité virtuelle. Ces casques permettent à l'utilisateur d'être à l'intérieur d'un monde qui n'existe pas, tout en ayant l'impression d'y être sans y être. Ainsi les tests demandant un éclairage particulier peuvent tous être réalisés dans des conditions parfaites. Les tests de mouvement permettant d'identifier la convergence oculaire peuvent aussi être réalisés dans un environnement sans variante. Quand un humain fait ce test,

il ne déplace pas toujours l'objet générant les stimuli visuels selon une trajectoire circulaire à une vitesse constante et reproductible.

En comparant l'intelligence artificielle et les humains dans l'application, des tests symptomatiques, une règle bien simple permet de décrire ce qu'un humain fait lors de la création et de l'application d'un test : l'humain résume et prend la version la plus simple, pouvant admettre des erreurs. L'intelligence artificielle de son côté, oui, simplifie le problème, mais cherche à minimiser l'erreur et non pas à avoir la représentation la plus simple. Un bon exemple serait la maturité d'une pomme, l'humain dirait qu'elle a besoin d'une certaine taille et d'une couleur. Un ordinateur pourrait plutôt analyser que si la taille est petite, mais qu'elle est très rouge elle est prête malgré qu'elle n'ait pas atteint sa taille, car il mettrait une pondération à chaque critère plutôt que de devoir absolument avoir chaque critère.

1.1 Objectifs de ce mémoire

Ce mémoire a plusieurs objectifs, le premier est de séparer les tests symptomatiques tels que le HGN (Horizontal gaze nystagmus/test de nystagmus horizontal) des distractions extérieures. Le second est de prouver qu'il est possible de réaliser les tests symptomatiques de détection des facultés affaiblies établis dans le DRE (Drug recognition expert) à l'intérieur d'un casque de réalité virtuelle. Les tests du DRE qui seront pris en compte dans cette recherche seront le test du nystagmus, de convergence et de dilatation oculaire. Le troisième objectif est de prouver qu'un algorithme de détection des patrons sous forme d'arbre peut être utilisé pour détecter l'état des facultés de la personne plutôt que d'utiliser des règles préétablies.

Ce mémoire sera séparé en plusieurs chapitres, commençant par les faits reconnus par les recherches des étudiants précédents ayant travaillé sur cette même problématique de détection automatique des facultés affaiblies. Suivi d'une brève description des algorithmes N-Gram et de leurs limitations qui seront utilisés pour représenter la connaissance extraite des tests symptomatiques implémentés pour la détection automatique des facultés affaiblies. Ensuite il sera question de l'algorithme de représentation en arbre et du dispositif de capture. Ce mémoire se terminera alors sur l'analyse des résultats et d'une preuve de concept de l'algorithme

d'identification des facultés. Le chapitre deux présentera les concepts qui ont mené à l'algorithme de représentation et les recherches effectuées dans des sources connues. Le troisième chapitre présentera la méthode utilisée dans cette recherche pour traiter les données. Ensuite, il sera question des résultats et de leurs analyses. Finalement ce mémoire se terminera avec un résumé des résultats et des ouvertures possibles ainsi que les points qui pourraient avoir été améliorés.

CHAPITRE 2

REVUE DE LITTÉRATURE

2.1 Recherches antérieures

Cette recherche est une suite des recherches de plusieurs étudiants s'étant échelonnées sur quelques années. Dans les sections qui suivent, un survol de l'évolution des prototypes développés pour la détection automatique des facultés affaiblies d'une personne.

2.1.1 Premier prototype

Le but principal de cette première recherche était d'automatiser le test du Nystagmus horizontal (HGN) [4]. Ceci permit d'éliminer l'interprétation humaine principalement au niveau de l'analyse des résultats du test HGN. Dans le cadre de cette recherche, le test HGN a été implémenté avec une caméra web permettant la capture des séquences d'images du mouvement de l'œil gauche en suivi oculaire horizontal. Ces séquences d'images sont ensuite analysées pour permettre de détecter la présence du HGN. Cependant, puisque le déplacement de l'objet générant le mouvement oculaire était encore fait par un humain, l'objet ne se déplaçait pas à une vitesse prévisible pour l'algorithme de détection du HGN. Un point fort de ce projet a été de présenter une preuve de concept sur la pertinence d'utiliser le nystagmus pour détecter les facultés affaiblies. Cela servira de base à tous les étudiants suivants ayant travaillé sur cette même thématique de recherche.



FIGURE 1 PROTOTYPE UTILISÉ PAR PIERRE DROLET [4]

2.1.2 Second prototype

Un second prototype (Figure 2) [5] a été développé à partir du premier prototype [4]. Le but principal était aussi d'automatiser le même test implémenté par Drolet [4]. Il a ajouté aussi un test de dilatation pour permettre la reconnaissance d'autres substances. En rendant le test complètement sans interaction humaine, il pouvait donc être considéré valide légalement. Il a construit alors la base d'un système autonome qui pourrait être utilisé sans avoir les problèmes liés à la formation du policier. Chaque catégorie de substance est abordée en détail dans le mémoire [5](p.25-33) afin de pouvoir traiter manuellement chaque catégorie de substance dans l'application. Il a créé le dispositif de capture autonome original. Ce dispositif a été basé sur un casque de soudure créant un environnement intérieur sombre, sans lumière externe. De plus une mentonnière a été ajoutée pour le confort et le positionnement stable de la tête. Il s'est servi d'une grande quantité de diodes lumineuses rouges afin de pouvoir produire l'illusion de mouvement à suivre. Le mouvement oculaire est alors généré automatiquement par une série de diodes électroluminescentes (DEL) rouges qui s'illuminent en séquence continue du centre vers l'extérieur. Ce nouveau dispositif permet de capturer les deux yeux en même temps.



FIGURE 2 PROTOTYPE UTILISÉ PAR PIER-OLIVIER HOUDE [5]

2.1.3 Quatrième prototype

Le quatrième prototype développé par Adam [6] a ajouté le test de convergence à la détection des facultés affaiblies. Il a pris le projet de Laperrière [7] qui avait remplacé les diodes rouges par un écran à cristaux liquides et en a ajusté les spécifications pour une meilleure détection. Il a ajouté de nombreux procédés au traitement du vidéo afin d'améliorer la qualité de la détection de la pupille et de l'iris. Les travaux d'Adam permettent de définir le procédé de traitement des vidéos (Figure 3) qui sera utilisé par les étudiants ayant travaillées depuis sur les projets subséquents.

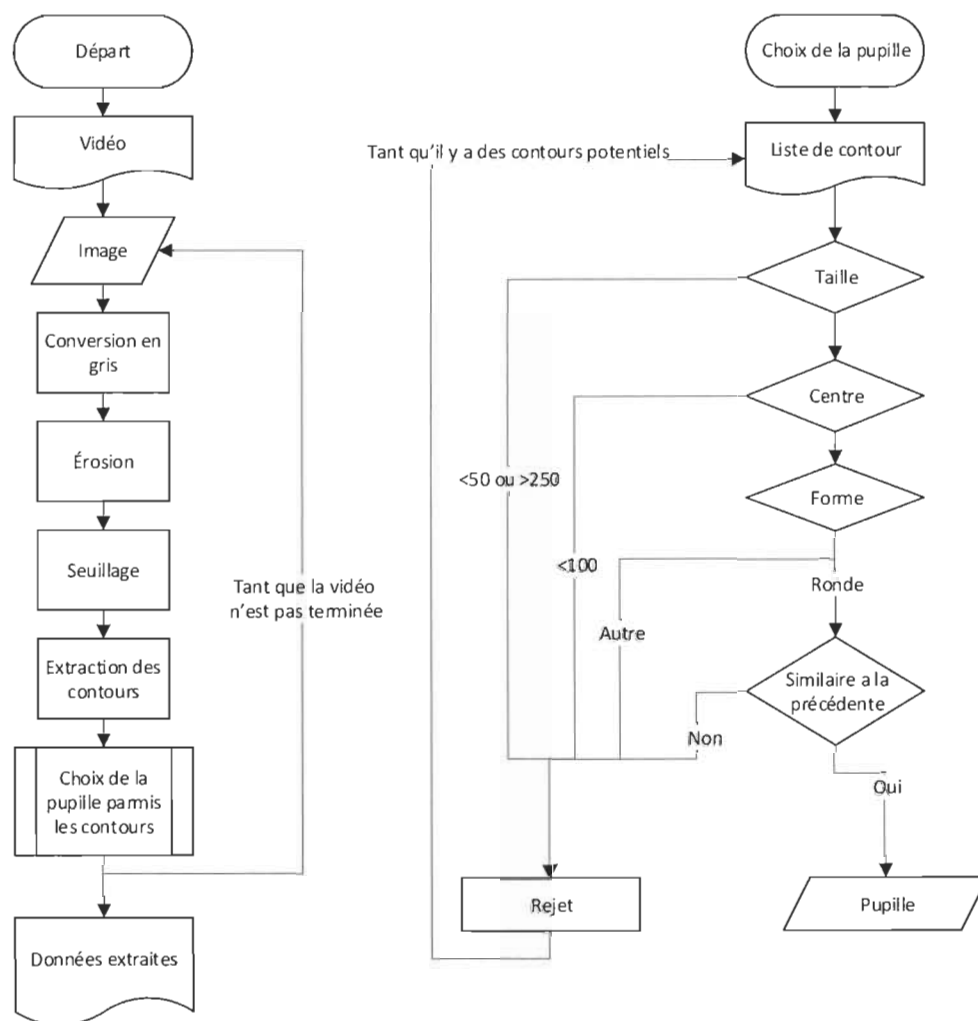


FIGURE 3 PROCESSUS EXTRACTION DES PUPILLES INTRODUIT PAR DOMINIC ADAM [6]

2.1.4 Premier prototype sur plateforme mobile

Les travaux de Lecours [8] ont porté sur l'adaptation du logiciel de détection des facultés affaiblies développé dans les recherches antérieures pour fonctionner sur une tablette Surface Pro de Microsoft. Dans un premier temps le but était de rendre le dispositif beaucoup plus portable. Il y est arrivé en utilisant une tablette ou un téléphone comme dispositif de capture plutôt qu'un ordinateur portable avec plusieurs de dispositifs externes. Son adaptation a rencontré de grands problèmes quant à l'éclairage ambiant pouvant rendre la détection des pupilles très difficile. De plus, la caméra de la tablette est de très grand angle ce qui induit de fortes déformations géométriques dans les images de pupilles.

2.2 Revues de littérature reliée au sujet de recherche

2.2.1 Anatomie de l'œil

Pour bien comprendre les tests symptomatiques basés sur le comportement oculaire à effectuer, il faut d'abord acquérir quelques connaissances sur l'œil humain. La présente recherche utilise la pupille et l'iris afin d'analyser les facultés d'un sujet.

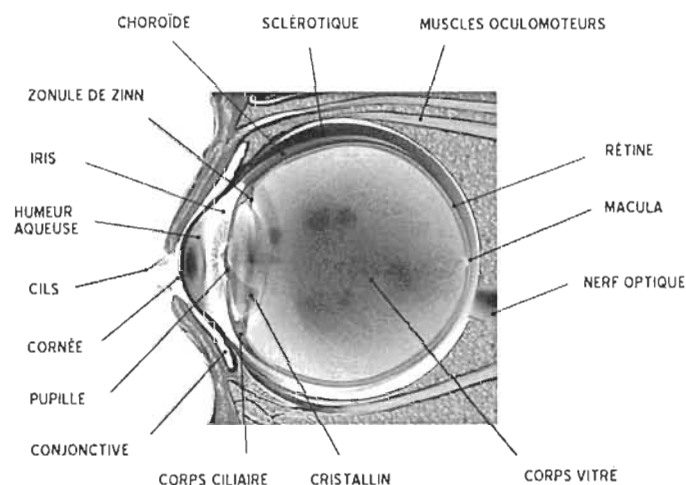


FIGURE 4 ANATOMIE DE L'ŒIL HUMAIN RÉCUPÉRÉ SUR GUIDE VUE [9]

Lorsqu'un humain a les facultés affaiblies, les muscles oculomoteurs (Figure 4, haut droit) et l'iris (Figure 4, gauche centre élevé) sont affectés, ce qui occasionne les changements des réactions normales qui peuvent être observés lors de certaines actions. Par exemple, lorsqu'on regarde vers l'extrême gauche, l'œil peut produire des saccades brusques et involontaires causées par le muscle oculomoteur, c'est ce qu'on appelle le Nystagmus horizontal.

2.2.2 Tests standards

Des tests standards, aussi appelés tests de sobriété standardisés, ont été développés pour identifier si la personne présente un nystagmus. Ces tests sont simples et peuvent être utilisés par les policiers au bord de la route afin d'évaluer l'état des facultés de conduite de la personne. Bien que le nystagmus se retrouve naturellement pour une très petite proportion de la population, sa présence indique généralement la consommation d'alcool ou de dépresseurs du système nerveux central. Le test qui est actuellement réalisé manuellement par les policiers pour évaluer l'état de la personne peut durer approximativement douze secondes. On présente un objet à la personne testés qui sera initialement de dix à quinze centimètres devant son nez. Ensuite on déplace cet objet à une vitesse constante lente de quinze à vingt degrés par seconde, soit alors d'une durée de quatre secondes environ, vers un angle de regard de 60 degrés vers l'extérieur. On tient l'objet à cet angle pendant quatre secondes avant de faire le mouvement inverse pour le retour vers le nez. Dans ce test, il est aussi possible de trouver plusieurs informations. Dans le cas d'intoxications à l'alcool, l'apparition du nystagmus avant l'angle de 45 degrés représente généralement le niveau d'intoxication. Un manque de suivi oculaire, ou un suivi irrégulier de l'objet peut aussi représenter une intoxication. La présence d'énormes saccades proches du regard latéral (45 degrés, voir Figure 6 droite) représente elle aussi une intoxication. Dans le diagramme présenté à la Figure 5, le cercle représente la tête de l'individu et le « N » représente son nez. Cette figure de représente que le test pour un œil, il faut alors faire le test sur l'autre œil en effectuant à même action de l'autre côté de la tête.

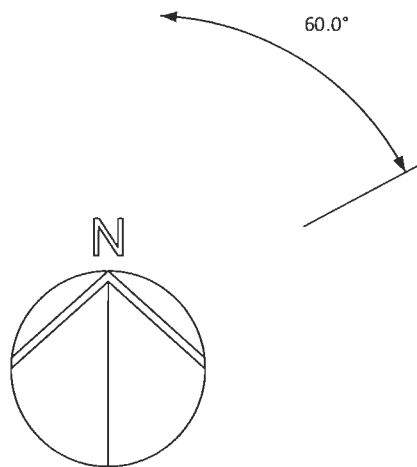


FIGURE 5 TEST DU NYSTAGMUS HORIZONTAL DROIT.



FIGURE 6 IMAGES CAPTURÉES PENDANT LE TEST DU NYSTAGMUS HORIZONTAL.

Normalement, l'iris se dilate ou se contracte par réflexe naturel pour adapter l'œil à son environnement lumineux. [9] Cela dit, on a l'impression que c'est la pupille qui change de taille, mais la pupille est plutôt le trou central de l'iris. Lorsque les facultés sont affaiblies, l'iris se dilate et se contracte à une vitesse différente, généralement plus lentement, ce qui peut causer des aveuglements sur la route. La Figure 7 permet d'observer le phénomène de dilatation et de contraction.

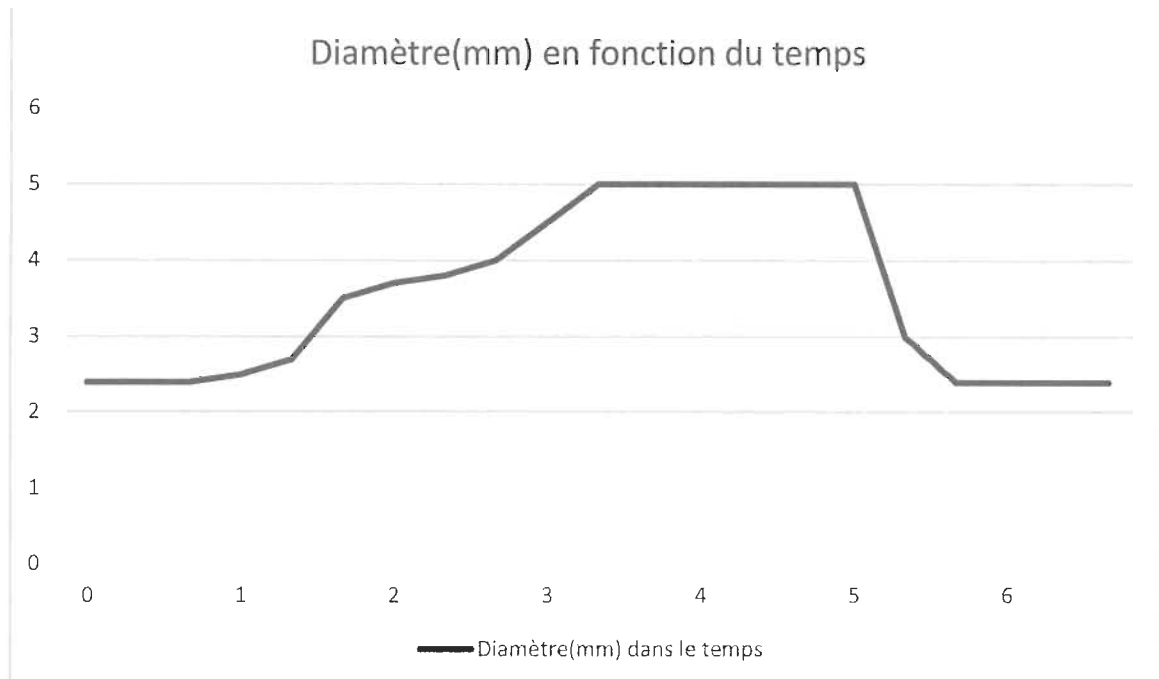


FIGURE 7 COURBE DE DILATATION/CONTRACTION

Un autre test standard pour l'identification des facultés affaiblies est celui de dilatation et de contraction oculaire. Ce test très simple à effectuer lorsque l'environnement nous le permet consiste à estimer le temps requis pour que l'œil adapte son ouverture à la lumière ambiante et à une lumière intense. Pour la réalisation de ce test, on place le sujet dans l'obscurité quasi totale afin d'identifier la taille de sa pupille au repos. Cette information est observable à la Figure 7 lorsque le diamètre augmente. La Figure 7 permet d'observer la variation de la dimension de la pupille. La dimension étant plus petite quand l'illumination est importante (Partie de gauche de la courbe) et plus grande quand l'illumination est faible (partie de droite de la courbe). Quand le test est effectué, il est possible d'observer une personne en portant le faisceau d'une lampe de poche allumée vers ses yeux pour ainsi avoir sa pupille aussi petite que possible (voir Figure 8 gauche). Cette même lampe de poche sera alors éteinte après quatre secondes ce qui permettra d'apercevoir une dilatation de la pupille (Voir Figure 8 droite). Le temps de dilatation ou de contraction peut alors être estimé. Un temps de contraction de la pupille de plus d'une seconde démontre les facultés affaiblies d'une personne, du moins pour l'alcool. Certaines autres substances affecteraient le test différemment et alors la différence par rapport à la courbe standard peut alors être évaluée.

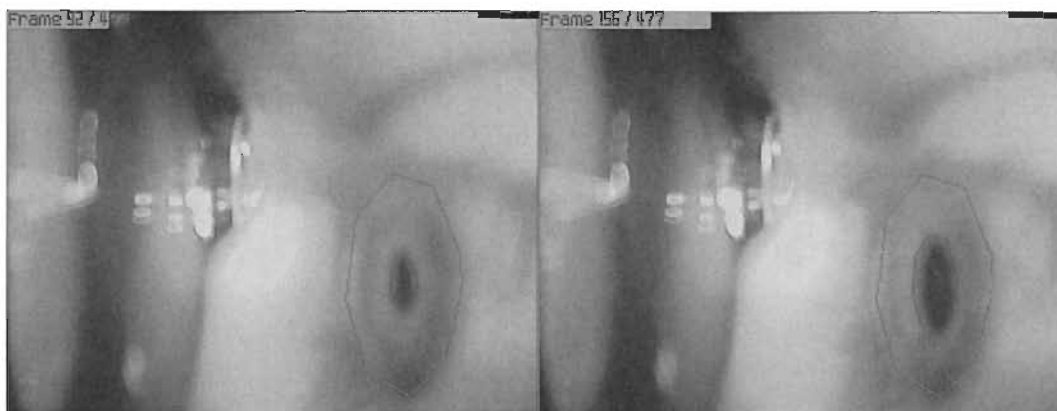


FIGURE 8 IMAGES CAPTURÉES LORS D'UN TEST DE DILATATION

Le troisième test standard qui sera abordé dans ce mémoire est celui de la convergence oculaire. Ce test sert à identifier si la personne est capable de suivre du regard un objet précis peu importe sa position. Ce test consiste à déplacer un objet de manière elliptique à convergence centrale à une vitesse constante afin d'identifier si l'œil peut le suivre de manière douce. À la fin du test l'objet devrait se retrouver sur le nez de la personne afin que les yeux puissent converger. Un manque de convergence à la fin ou un suivi irrégulier de l'objet peut identifier une intoxication quelconque. La Figure 9 permet d'abord d'observer la trajectoire de suivi oculaire requis par le test de convergence. La Figure 9 permet aussi d'observer quelques exemples d'images de suivi oculaire pour différentes positions de la trajectoire circulaire générée.

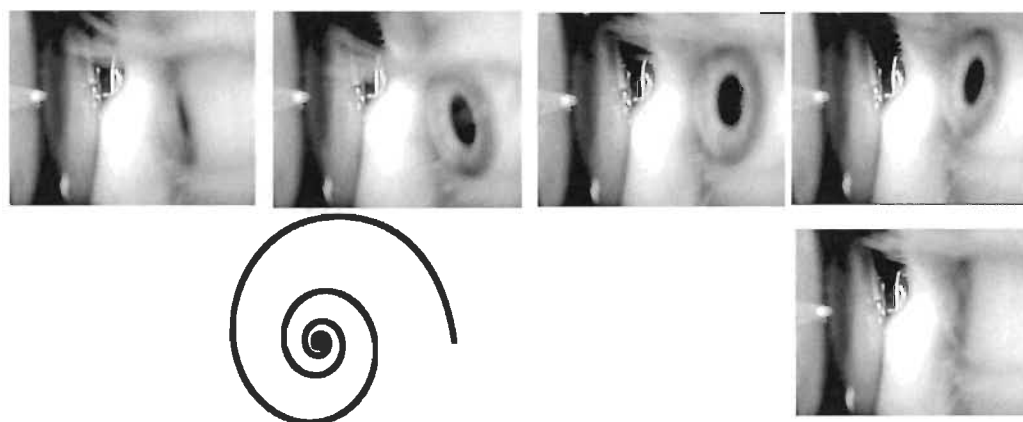


FIGURE 9 IMAGES CAPTURÉES LORS D'UN TEST DE CONVERGENCE ET DÉPLACEMENT DU STIMULUS

2.2.3 Analyse des yeux par ordinateur

Lors de l'analyse des vidéos de mouvement oculaire par ordinateur, il faut faire attention de bien identifier les bonnes parties de l'œil, la pupille et l'iris sont les données désirées. Pour le test de contraction et de dilatation, on regarde le ratio de pupille sur iris afin d'en extraire la taille. Pour celui du nystagmus horizontal, on peut identifier l'angle de visée de la personne et identifier les saccades brusques (Nystagmus). Il est possible d'extraire l'angle de plusieurs façons, deux seront abordées dans ce mémoire. Il faut d'abord faire l'extraction des régions d'intérêts (ROI) dans le champ de vision des caméras pour ainsi distinguer les régions utiles et nuisibles. Les régions nuisibles sont la cornée qui pourrait refléter une source d'illumination et donc créer des interférences, les cils et l'humeur aqueuse.

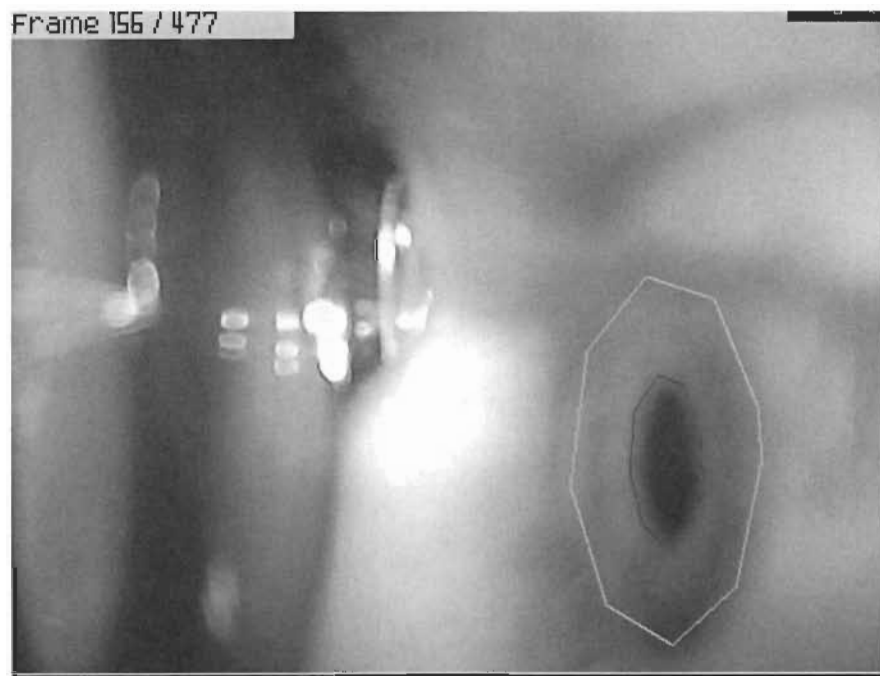


FIGURE 10 IMAGE D'UN ŒIL

Différentes techniques existent pour l'extraction des données découlant du ROI de l'œil (pupille/Iris). En ce qui a trait à l'identification de la taille de la pupille, il est possible de l'identifier dans un premier temps par un rapport pixel sur millimètre. Un problème avec cette technique d'extraction est que la distance de la caméra doit être constante, ce qui dans notre cas n'est pas

toujours le cas. Une deuxième technique est d'identifier l'iris, puisqu'il est de taille connue et à la même distance que la pupille dans l'image, il est possible d'en calculer un ratio afin d'obtenir la taille en millimètres de celle-ci.

En ce qui a trait à l'identification de l'angle de visée, il est possible de calculer celui-ci d'après un ratio largeur sur hauteur de la pupille. Bien que cela dépende de l'angle de la caméra, il est possible de trouver la formule en effectuant une calibration sur les extrémités connues. Par exemple si on sait que la caméra se trouve à 45 degrés par rapport à un regard vers l'avant et que l'œil regarde initialement principalement en avant (0 degrés), il est possible de faire une extrapolation linéaire ou cubique afin d'obtenir l'ensemble des angles entre ces deux points. Une autre technique pour l'identification de l'angle est une calibration selon le centre de la pupille. Puisque le ratio peut être altéré si jamais un objet tel qu'un cil cache une partie de la pupille, le centre risque de ne pas être affecté. En utilisant celui-ci plutôt que le ratio, il est aussi possible d'obtenir une équation polynomiale. Cependant en utilisant le centre dans notre formule, nous devenons dépendants de la résolution de la caméra et de son positionnement.

2.2.4 Alcoolémie

Lors de la consommation d'alcool, le corps humain réagit de plusieurs façons face à ce liquide. L'alcool peut augmenter le temps de réaction, rendre somnolent, nuire à la vision et rendre les choix moins sensés. C'est pour cela que la consommation d'alcool est limitée lorsqu'on conduit ou effectue des tâches complexes telle qu'opérer de la machinerie. En conduite automobile la quantité maximale d'alcool est définie par le Code la sécurité routière [10]. Il existe plusieurs moyens d'identifier le taux d'alcool qu'à consommer un humain, soient des tests sanguins ou des tests d'haleine. Les tests sanguins sont plutôt rares dû au temps d'obtention des résultats. Les tests d'haleine sont moins précis, mais ont une réponse de l'ordre de la minute. La mesure pour identifier le taux d'alcool dans le sang est la masse d'alcool en milligramme(mg) par une centaine de millilitres de sang, soit mg%. On parle bien de milligramme pour cent millilitres de sang étant donné que la quantité d'alcool n'a pas besoin de dépasser une centaine de milligrammes pour rendre la personne inapte à conduire. Selon le C.S.R. [10] il faut un taux maximal de 80 mg% pour être inapte à la conduite automobile. Un taux moins élevé d'alcool dans le sang ne démontre cependant pas pour autant le fait que la personne est apte à conduire.

Chaque personne répondant différemment à l'alcool, une valeur absolue pour être considéré apte à la conduite n'est donc pas identifiable. Cependant, au-delà de 80 mg%, il ne devrait pas y avoir d'humain apte à conduire. Il est alors question d'un niveau acceptable d'intoxication au niveau légal et non au niveau de l'aptitude à conduire.

Étant donné l'existence de test standard d'identification des facultés affaiblies, il est possible qu'un automobiliste soit reconnu coupable de conduite avec les facultés affaiblies malgré un taux d'intoxication inférieur à ce qui est inscrit dans la loi. Il suffit à cette personne d'échouer un certain nombre de ces tests standardisés pour être déclarés coupables, et cela, peu importe son niveau d'intoxication.

2.2.5 Autres drogues

Ce n'est pas uniquement l'alcool qui affecte les facultés d'un humain à effectuer des tâches complexes comme la conduite automobile. Bien des drogues telles que le Cannabis ou la Cocaïne rendent rapidement l'humain hors d'état d'effectuer ses tâches. Il n'est pas du rôle de ce mémoire d'identifier les taux légaux de chaque substance, mais il était important de les considérer, car le système développé n'est pas limité à l'alcool, car il est aussi calibré pour ces différentes substances. Un résumé des effets sur les yeux est inclus pour identifier les effets que le système serait entraîné à reconnaître avec une bonne base d'apprentissage. Cette liste est tirée de American Addiction Centers [11] et [12].

- L'alcool peut créer une vision floue ou double qui peut être analysée dans les tests de suivi oculaire.
- L'amphétamine peut causer du nystagmus, une vision double et une pupille de taille anormale.
- La cocaïne dilate la pupille pouvant être analysée par une vitesse de contraction/dilatation différente en changement d'illumination.
- L'héroïne contracte la pupille et affecte le taux de dilatation qui dans ce contexte est nul.
- Le Cannabis rend le blanc de l'œil injecté de sang, une caméra couleur pourrait détecter cette couleur irrégulière.
- Les narcotiques peuvent empêcher la réponse de la pupille à la lumière, celle-ci reste petite pour cette condition d'illumination. (Voir héroïne)

2.2.6 Traitement d'image avec Emgu.CV

Afin de traiter les images électroniquement, il est nécessaire d'utiliser des algorithmes. EMGU.CV est une librairie utilisant le Framework « .NET » basé sur la librairie « Open CV » [13]. Le but de ces librairies est de rendre accessible à tous, le traitement d'image. Il existe plusieurs dizaines de formats d'image, beaucoup de marques et modèles de caméras, cependant, Open CV en fait abstraction pour un traitement uniforme. La raison principale de l'utilisation de cette librairie est qu'elle permet d'acquérir les images d'une caméra. Une autre raison de l'importance d'une librairie telle qu'Open CV est l'optimisation des algorithmes et l'utilisation de la carte graphique pour le traitement de ceux-ci.

Mes prédécesseurs ont tous utilisé Open CV, leur processus de traitement peut être défini par des opérations de base. [6]

1. Obtention de l'image par un objet « VideoCapture »
2. Transfert de l'image en format niveau de gris
3. Application d'érosion et de lissage
4. Application de seuils(segmentation)
5. Détection des contours (Un contour est une délimitation entre deux segments d'image définis par un changement de couleur brusque)
6. Choix du meilleur contour pour l'information désirée
7. Extraction de la donnée pour le test courant

Ce procédé possède certaines failles qui seront corrigées dans cette recherche. Premièrement, les seuils sont dépendants de l'éclairage ambiant et de la caméra utilisée. Pour faire une meilleure abstraction de l'éclairage, il serait judicieux de plutôt utiliser une proportion de l'histogramme des niveaux de gris plutôt qu'un seuil absolu. Un autre aspect pouvant être considéré comme une faille serait l'extraction des données pour le test courant. Il serait plus judicieux de toujours extraire l'ensemble des données et de n'utiliser que les données nécessaires étant donné que le traitement d'une vidéo est long.

Certains concepts ont été discutés dans le processus des prédécesseurs, ils seront maintenant décrits avec plus de détails. D'abord, le processus d'érosion (opération 3 à la Figure 3) est un processus permettant de retirer une partie d'une image telle qu'une ligne. Le processus de lissage est un terme générique pour l'application d'un flou gaussien. Ces deux processus sont en fait des filtres de convolution de l'image en appliquant des noyaux différents. Étant donné que la librairie fait abstraction de l'implémentation il n'est pas possible de voir le noyau utilisé cependant on peut en supposer qu'une équipe a optimisé ces fonctions pour les meilleurs résultats possibles.

Bien qu'Open CV soit une librairie multi-plateforme, son utilisation dans le .NET Framework sur Windows est dépendante d'une autre librairie (System.Drawing/ GDI+). Il n'est donc pas possible de porter le traitement effectué sur une autre plateforme supportant ce framework tel que Android.

2.2.7 Traitement d'image avec ImageProcessing

Une librairie que j'ai développée (ImageProcessing) en dehors du cadre de cette recherche est de source libre accessible sur GitHub afin de traiter les images. [14] Elle reprend le principe d'Open CV en boîte noire sur le traitement, tout en offrant plus de contrôle à l'utilisateur sur ce qui se produit. Les grandes différences entre ces deux librairies est le traitement. Dans Open CV chaque traitement crée une nouvelle image, de plus chaque sous-section de l'image est aussi une nouvelle image plutôt que d'être une simple passerelle d'accès à ces données. Dans ImageProcessing, la majorité des traitements sont effectués en modifiant l'image plutôt que d'en créer de nouvelles, ce qui permet d'utiliser le programme sur un ordinateur ayant moins de mémoire vive. Lorsqu'on désire accéder à une région d'intérêt, on spécifie cette région et on obtient une passerelle, soit un moyen d'accéder à l'information qui n'est pas une copie.

En ce qui a trait au traitement de la vidéo, cette librairie ne prend pas en charge les vidéos, mais accepte volontiers tous les formats d'image statique et surtout les Bitmaps. Étant donné que Open CV gère les vidéos, il a été convenu d'extraire les images avec Open CV et ensuite les transférer dans ImageProcessing simplement en spécifiant l'endroit où se trouve l'image en

mémoire (celle réservée par Open CV). Par la suite la majorité des traitements possibles en Open CV sont disponibles, ceux qui ne le sont pas peuvent facilement être ignorés et obtenir un très bon résultat de détection.

Pour continuer le parallèle d'Open CV, avec ImageProcessing, lors de la détection des contours, il est possible de spécifier le critère de sélection des contours. Open CV détecte automatiquement les grands changements de luminosité dans les contours pour nous et parfois il sélectionne le contour qui n'est pas nécessairement le contour désiré. Cette spécification du critère permet aussi d'éliminer l'étape du seuillage. (Voir Figure 11) Un exemple de critère serait que le seuil de luminosité est supérieur à 50% afin de sélectionner les sections blanches de l'image.

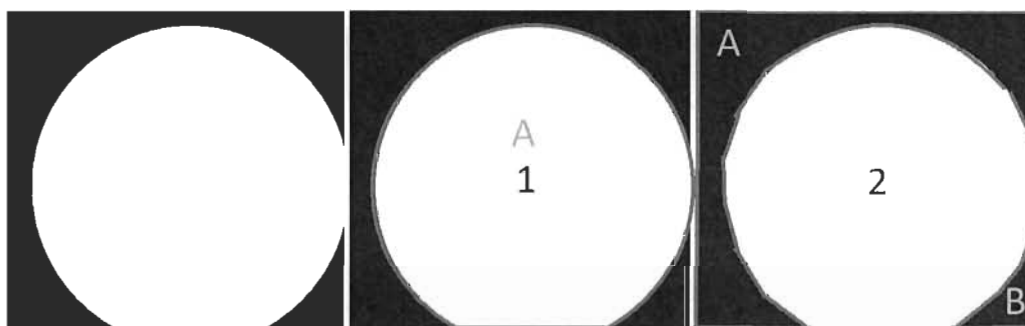


FIGURE 11 SEGMENTATION EN OPEN CV

À GAUCHE, IMAGE DE RÉFÉRENCE, AU CENTRE IMAGE BIEN SÉLECTIONNÉE POUR LES CONTOURS, À DROITE SÉLECTION INVERSE ET NON DÉSIRÉE

En ce qui a trait au seuillage, il est possible dans ImageProcessing de faire un seuillage intelligent, soit d'identifier un pourcentage des couleurs à conserver plutôt qu'un seuil établi d'avance, qui a pour but d'abstraire le problème d'éclairage. Un bon moyen pour établir le seuil est de faire un algorithme K-MEANS [15]. Cela permet alors d'obtenir un certain nombre de catégories de couleurs. La librairie permet alors de réutiliser le résultat dans les images suivantes plutôt que de réappliquer le processus. Comme énoncé précédemment, le seuillage ne crée pas une nouvelle image comme dans Open CV et il est même possible qu'il ne la modifie pas. L'algorithme de détection des contours, ou plus précisément, dans cette librairie, des zones de succès, n'altère pas l'image d'origine en identifiant les zones correspondant à un critère de sélection. Ce critère est généralement un seuil sur la couleur, mais peut être défini autrement.

2.3 Revue de la littérature reliée directement aux nouvelles spécialités de cette recherche

2.3.1 N-Gram

Un N-Gram est une suite contiguë de symboles dans une séquence plus longue [16]. Cette méthode de classification automatique qui sera utilisé dans cette recherche par de multiples auteurs en obtenant des résultats de très bonne classification. [17] [18] [19] [20] [21]

Pour évaluer la similitude de deux séquences, il est possible de compter le nombre de N-Gram communs sur le nombre total de N-Gram possible dans ces séquences. Ces N-Gram sont généralement utilisés dans le traitement des langues naturelles (TLN). Une grande quantité de N-Gram communs indique un mot similaire. [18]

Mot = Précéder

N-Grams possibles :

P,r,é,c,d,e,r,Pr,ré,éc,cé,éd,de,er,Pré,réc,écé,céd,éde,der,Préc,récé,écéd,céde,éder,Précé,récéd,écéde,céder,Précéd,récéde,écéder,Précéde,récéder,Précéder

VS

Mot = Précédemment

N-Grams possibles :

P,r,é,c,d,e,m,n,t,Pr,ré,éc,cé,éd,de,em,mm,me,en,nt,Pré,réc,écé,céd,éde,dem,emm,mme,men,Préc,récé,écéd,céde,édem,demm,emme,mnen,ment,Précé,récéd,écéde,cédem,édemm,demme,emmen,mment,Précéd,récéde,écédem,cédemm,édemme,demmen,emment,Précéde,récédem,écédemm,cédemme,édemmen,demment,Précédem,récédemm,écédemme,cédemmen,édemment,Précédemm,récédemme,écédemmen,cédemment,Précédemme,récédemmen,Précédemmen,Précédemment

FIGURE 12 EXEMPLE DE N-GRAM EN TLN

Pour une séquence donnée, la méthode principale d'identification des N-Gram est une série de trois boucles imbriquées (voir Figure 13). Il est possible d'utiliser les N-Gram de longueur fixe, cependant, en utilisant l'ensemble des longueurs possibles de N-Gram, plus d'information est extraite.

```

Knowledge = {0}

For begin from 1 to Length
    For length from 1 to (Length-Begin)
        N-Gram = source [begin, length]
        For N-GramK in Knowledge
            If N-Gram == N-GramK
                N-GramK ++
        If N-Gram not in Knowledge
            Knowledge += N-Gram

```

FIGURE 13 ALGORITHME N-GRAM

La principale raison pour utiliser un N-Gram est qu'il n'est pas très sensible au bruit. Puisqu'une séquence est composée de plusieurs sous-séquences qui se superposent, la méthode des N-Gram permet d'avoir une bonne classification même si un bruit ajoute une sous-séquence imprévue. [18]. Cette méthode est alors très performante dans le cas où la source de donnée n'est pas très fiable ou est portée au bruit telle qu'une détection par ordinateur d'un texte ou dans la présente recherche [22], de la localisation des pupilles dans un vidéo. Puisque la source de donnée peut être elle-même bruitée et que le processus d'identification apporte lui aussi son appart de bruit, une tolérance au bruit est nécessaire. La fréquence d'un N-Gram dans un texte permet de catégoriser ce texte. Il y a toujours les N-Gram reliés à la langue courante. Mais principalement la différence de fréquence d'un N-Gram permet de calculer une différence entre plusieurs catégories. Le sujet du texte est alors représenté par le sujet ayant la plus petite distance parmi tous les sujets possibles.

Toujours selon Cavnar [18], les N-Gram ont un avantage sur les autres types d'algorithmes, si ces algorithmes sont appliqués sur des textes, ils permettent d'obtenir implicitement l'élagage (word stemming) des mots. Par exemple, les mots « aimer » et « aimé » deviendront assez semblables pour être reconnus dus à la grande quantité de sous séquences similaires.

Le processus que Cavnar utilise dans son article de 1994 a de grands avantages quant à la détection des catégories. Voici un résumé de son processus :

1. Compter les N-Gram d'une source et en extraire un profil
2. Comparer le profil a l'ensemble des catégories apprises, chaque catégorie possède un profil appris.
3. Trouver la distance par rapport à chaque catégorie.
4. Choisir la catégorie en conséquence de la distance la plus courte.

La dernière étape peut être omise pour permettre d'avoir de l'information sur la distance entre les premiers choix si la catégorie est ambiguë.

2.3.2 Représentation des N-Gram en mémoire

Étant donnée la taille importante de l'information extraite des N-Gram, des moyens doivent être pris pour contenir l'information en mémoire de l'ordinateur qui peut être limitée (mémoire vive) ou lente (disque). « Décoder une simple phrase peut déclencher des centaines de milliers de requêtes au modèle, alors la vitesse est vraiment importante » [23] [Traduction libre]. Un moyen de représenter l'information de manière efficace en accès et petite en espace est de stocker nos N-Gram comme des nœuds dans des arbres. Un nœud représente un mot qui suit un contexte. Le contexte est alors la série de mots qui le précède. Récursivement, il est alors possible d'obtenir la séquence complète en partant du nœud le plus profond vers le nœud le moins profond, il ne suffira que d'inverser le résultat. La notion d'arbre peut alors être enregistrée dans des fichiers ou être en mémoire vive. Il est possible d'utiliser des objets avec des références dans un langage-objet ou simplement des indices dans un tableau (voir Figure 15) ou plusieurs tableaux (voir Figure 14). La notion d'objet permet de garder une grande quantité d'information

dans la mémoire vive de l'ordinateur effectuant le calcul, tandis qu'enregistrer les informations sur le disque ralentit beaucoup le calcul, mais permet une connaissance beaucoup plus grande.

Premier élément		Deuxième élément			Troisième élément		
ID	Mot	Précédent	ID	Mot	Précédent	ID	Mot
1	Le	1	1	Chat	1	1	Brun
2	Ce	1	2	Chien	1	2	Roux
3	Il	...	3	Oiseau	1	3	Noir
4	Nous	...	4	Cheval	2	4	Brun
5	Aimer	...	5	...	2	5	Roux

FIGURE 14 REPRÉSENTATIONS DE N-GRAM EN PLUSIEURS TABLEAUX

Élément		
Précédent	ID	Mot
0 (RACINE IMPLICITE)	1	Le
1	2	Chat
2	3	Brun
2	4	Roux
2	5	Noir
1	6	Chien

FIGURE 15 REPRÉSENTATIONS DE N-GRAM EN TABLEAU

2.4 Algorithme de représentation des résultats en arbre et comparaison

2.4.1 Introduction à l'algorithme en arbre

Un algorithme de classification des séquences d'information est nécessaire afin d'identifier les facultés d'un individu à travers un vidéo. Un des algorithmes intéressants est l'algorithme en arbre. Cet algorithme a été initialement développé lors de mes cours du baccalauréat pour l'identification du sujet d'un texte, mais n'ont pas été publiés en dehors des travaux de ces cours. Une brève publication de cet algorithme sera alors incluse dans ce mémoire pour l'introduire avant de l'appliquer au sujet de cette recherche. Elle est basée sur la notion des N-Gram utilisés par plusieurs auteurs, donc principalement Pauls, Kleins [23] et Cavnar [18]. Cet algorithme, indépendant de la langue, est alors utilisable dans autre chose que des langues telles que des signaux. Ce sera l'analyse par signaux qui sera utilisée pour l'appliquer à la détection automatique des facultés affaiblies. Un signal généré par un comportement humain, soit le nystagmus et tout mouvements de l'œil devraient se comporter de manière naturelle, soit de manière semblable à une langue écrite.

2.4.2 Résumé

Dans la littérature, de multiples textes existent et pourraient bénéficier d'une indexation ou d'un classement. Il existe de multiples algorithmes pour identifier le sujet d'un texte, mais ces algorithmes sont basés sur des critères linguistiques et sont difficiles à implémenter. Les méthodes statistiques actuelles ne permettent pas d'obtenir de bons résultats rapidement, car les filtres de mots ne peuvent pas s'appliquer au contexte. Il faut donc penser à un algorithme qui s'appliquera sur les suites de mots et non uniquement aux mots. Cependant les suites de mots dépendent des conjugaisons de verbes. Il faudra alors passer un filtre de prétraitement pour mettre les verbes à l'infinitif et tous les noms et adjectif au masculin singulier. L'uniformité est la seule option pour réussir une bonne détection du sujet d'un texte. Ensuite, nous enlèverons les mots communs (ex : le, la, de, du, ...). Ces mots peuvent être identifiés par un humain ou par analyse statistique de texte de multiples sujets, cette étape est optionnelle. Si tout le jeu de donnée provient d'un même sujet, il pourrait être difficile de distinguer les mots communs de la langue de ceux du sujet (ex en santé le mot santé). L'algorithme analyse le texte en créant un arbre des suites de mots avec plusieurs têtes d'écriture. Cet algorithme est basé sur la notion de

N-Gram, mais en ayant plusieurs avantages comparés à ceux-ci. La section 2.4.6 énumérera ces avantages en expliquant les différences de l'algorithme comparé aux N-Gram sans arbre.

2.4.3 Problématique

Lors de la recherche en traitement de langue naturelle, personne n'arrive à un consensus sur quelle information cibler pour extraire cette information. Certains avancent que le mot seul est suffisant, mais ces algorithmes ont fait leurs preuves et ne sont pas très précis [21]. D'autres affirment qu'une fenêtre autour du mot doit être utilisée pour identifier sa sémantique, cependant encore une fois les preuves ont été faites que ces algorithmes ne donnaient pas les meilleurs résultats [24]. La dernière option est une phrase complète, qui véhicule beaucoup d'information et qui donne actuellement les meilleurs résultats [24]. Se basant sur le principe des suites de mots, nous allons identifier les suites à l'intérieur d'une phrase pour ne pas à avoir la combinaison de toutes les fenêtres autour de notre mot comme inspiré par [17]. Nous créons un algorithme qui va extraire ces suites et les comparer pour obtenir un taux de similitude. Cet algorithme fonctionne en un passage sur le texte pouvant alors aussi s'appliquer aux textes non reproductibles tel que les informations qui viennent en temps réel et non d'un enregistrement.

2.4.4 Méthodologie (Les étapes)

2.4.4.1 Analyse du texte et représentation en arbre

Nous explorerons le texte sous forme d'arbre à racine implicite. D'abord, nous plaçons une tête d'écriture à la racine implicite. À chaque mot lu, nous demandons aux têtes d'écritures d'écrire le mot, ensuite les têtes se déplacent sur le nouveau mot et une nouvelle tête est ajoutée à la racine implicite. La Figure 16 montre les étapes de création d'un arbre de N-Gram. Les mots seront ajoutés un par un, soit le mot A, suivi du mot B, A et D.

Pour les mots A B A D

☆ = tête d'écriture

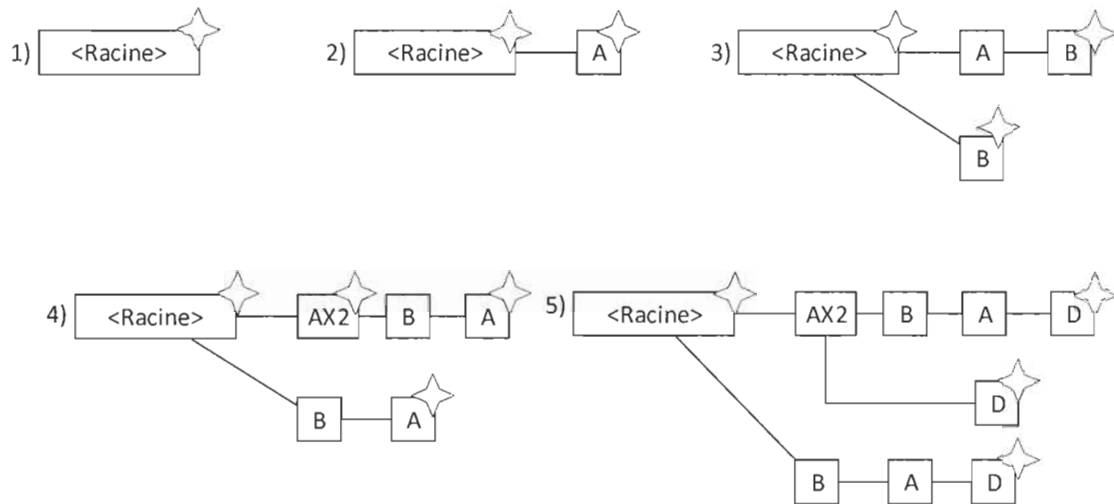


FIGURE 16 CRÉATION D'ARBRES SIMPLE

Voici en détail ce qui se produit pour chaque mot lu et ajouté à l'arbre. La tête regarde les successeurs du nœud actuel seulement pour choisir l'action à exécuter : soit insérer un nouveau successeur ou augmenter le nombre d'occurrences d'un successeur existant. À la fin de l'écriture, la tête se déplace sur ce successeur. Le nombre de têtes grandit avec la taille du texte, chaque mot ajoutant une tête.

Pour contrer ce phénomène, il faudrait alors insérer des points de coupes qui seront expliqués plus loin dans le texte. Un point de coupe étant le fait d'éliminer toutes têtes pour n'en conserver qu'une nouvelle à la racine.

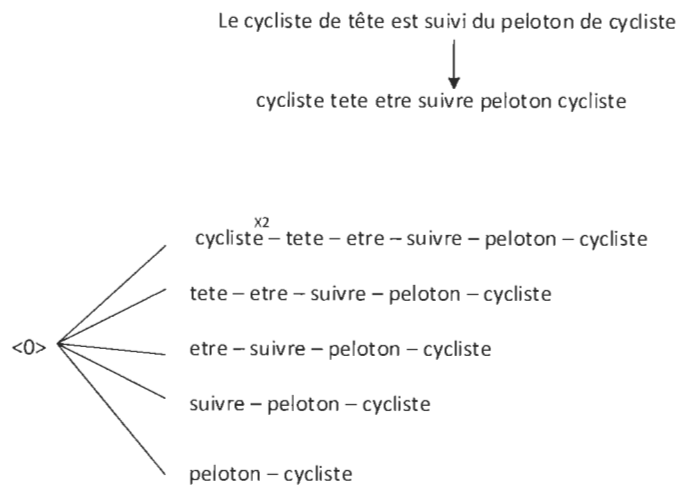


FIGURE 17 REPRÉSENTATIONS EN ARBRE D'UNE PHRASE

Ainsi, pour donner suite à la création de l'arbre sous cette méthode, il est possible de s'apercevoir que chaque branche, ou branche partielle représente un N-Gram (Figure 17). Il a donc été possible en un seul passage et en limitant le nombre de comparaisons aux nœuds seulement plutôt que l'ensemble complet de construire l'ensemble des N-Gram de toutes les longueurs possibles. Il aurait été possible en connaissant une longueur désirée de N-Gram de couper les têtes d'écriture pour toutes les branches trop profondes. Ensuite il aurait été nécessaire de couper les branches trop courtes pour la longueur de N-Gram désirée.

Meilleur des cas

Dans le cas d'une phrase contenant que le même mot, nous nous retrouverons avec N nœuds de ce mot et N têtes d'écriture qui ne font qu'une action par mot donc l'algorithme fonctionne en $O(2N^2)$. L'algorithme sortira alors un arbre linéaire représentant ce mot comme par exemple :

AxN	AxN-1	AxN-2	AxN-3	...	Ax5	Ax4	Ax3	Ax2	Ax1
-----	-------	-------	-------	-----	-----	-----	-----	-----	-----

Chaque mot lu incrémentera de 1 chaque élément de la liste(N) et ajoutera un nœud à la fin (1).

Puisque chaque itération fera N actions et qu'il y aura N itérations, chaque action comprend 1 comparaison et un ajout (Figure 18), on parle alors d'une complexité de $O(2N^2)$. La version gourmande des suites pour un même texte est représentée par N itérations première (chaque début), N itérations seconde (chaque longueur) et chaque ajout est succinct à N comparaisons, notre meilleur des cas représente alors un avantage sur le $O(N^3)$ qui ne réduit pas dans ce cas.



FIGURE 18 MEILLEUR DES CAS - ARBRE

Cas moyen

Prenons le cas d'une phrase contenant N mots différents, on parle toujours de lecture par mot donc N , chaque lecture crée un ajout N , mais nous savons qu'à la racine il y aura N comparaisons (Figure 19) et chacun des autres nœuds aura 1 comparaison, on parle alors d'une complexité de $O(2N^2 + N - 1)$.

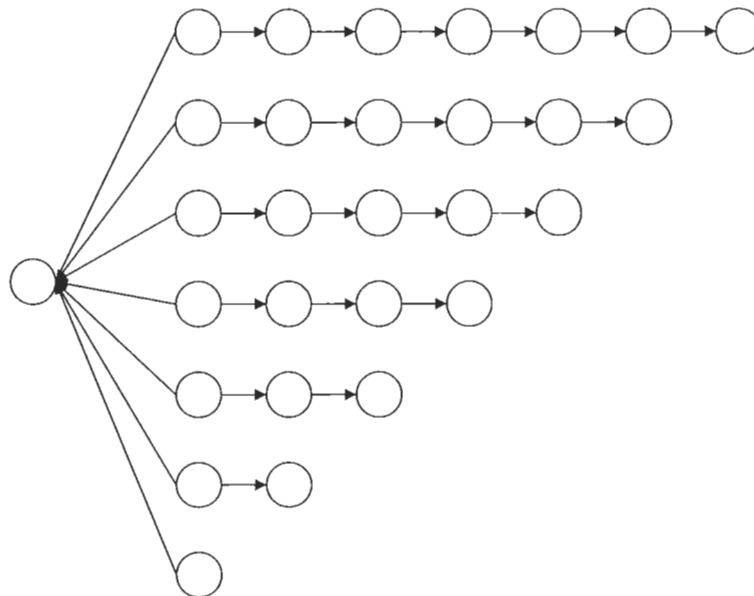


FIGURE 19 CAS MOYEN - ARBRE

Pire des cas

Prenons une phrase contenant une suite d'un mot alterné avec K autres mots, par exemple :

A B A C A D A

Cette phrase est similaire au cas moyen, car il n'y a pas une grande différence sémantique. On parle ici de moins de nœuds dû à la répétition de A, mais chaque itération aura deux fois plus de comparaisons dû au fait que A est répété et qu'il contient beaucoup de successeurs (Figure 20) comparativement au cas moyen ou chaque nœud n'en a qu'un. La complexité étant alors de $O(3N^2 + N - 1)$.

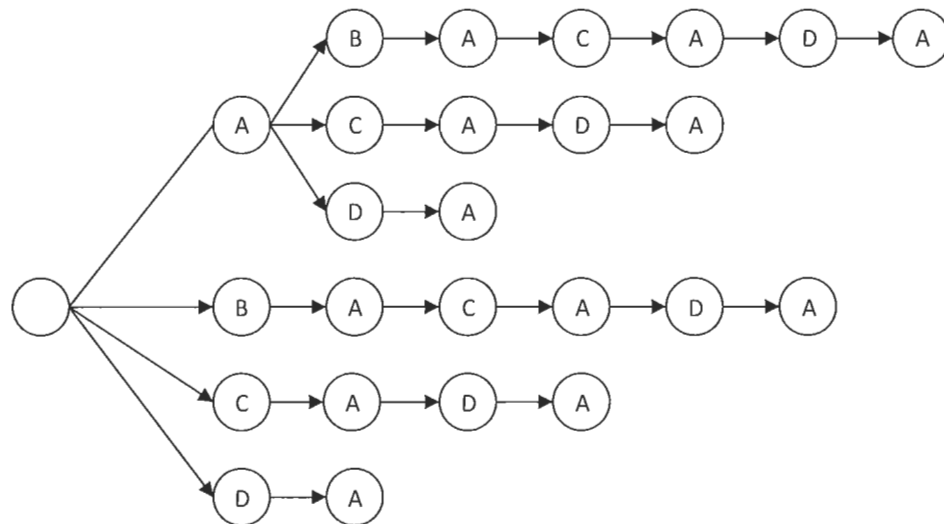


FIGURE 20 PIRE DES CAS - ARBRE

Comparaison avec la version N-Gram brute

Nous avons précédemment vu que notre meilleur des cas a une complexité de l'ordre N et le pire d'ordre N^2 et que l'exploration brute est d'ordre N^3 peu importe la phrase, cependant, nous n'avons pas considéré la différence fondamentale d'une comparaison, dans notre algorithme une comparaison étant faite sur un mot seulement d'une longueur de M lettres. Dans le cas de l'exploration brute, une comparaison se fait sur chaque mot de la suite, soit la longueur d'une phrase L maximale et 1 minimale ($L/2$) et M lettres par mot donne $L/2 + M$ pour chaque comparaison rendant l'algorithme brut beaucoup plus long pour une même comparaison. Un

autre avantage de l'approche par arbre est que chaque mot n'est lu qu'une fois plutôt que de resegmenter le texte à chaque itération. Si la segmentation du texte est lourde, l'algorithme en arbre permet de la minimiser.

2.4.4.2 Apprentissage/Addition

L'apprentissage de sujet se fait un peu sur le même principe que l'exploration, d'abord, il faut explorer le texte et générer l'arbre, ensuite il faut savoir à quel sujet appartient le texte, il peut alors y avoir plusieurs catégories. Dans une base de sujets, on vérifie premièrement que le sujet existe, s'il existe on additionne les arbres pour former un arbre plus grand et on l'ajoute dans la banque de sujets, si le sujet n'existe pas on prend un arbre vide lors de l'addition.

L'addition d'arbres ($A+B$) se fait de manière B vers A, soit pour chaque nœud de B, s'il existe dans A on augmente un compteur, initialement à zéro étant associé à ce nœud, sinon on ajoute le nœud à A. L'on applique alors cette règle récursivement jusqu'à ce que l'arbre complet soit additionné et donne un résultat C. (Voir Figure 21)

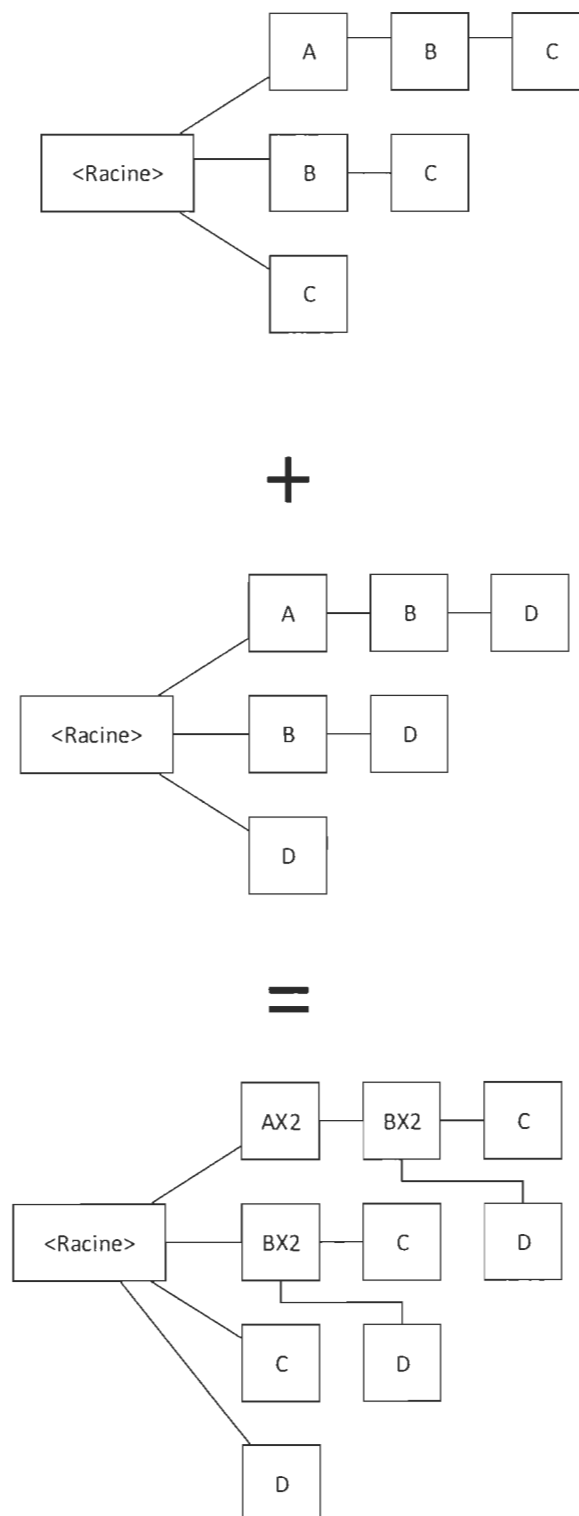
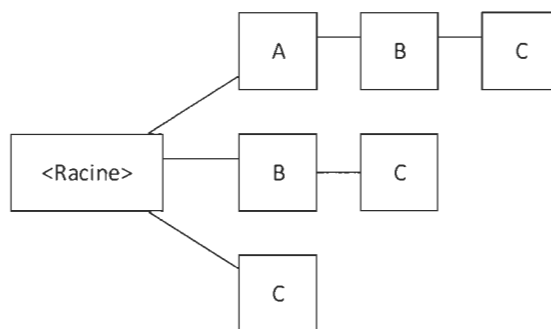


FIGURE 21 ADDITIONS D'ARBRE

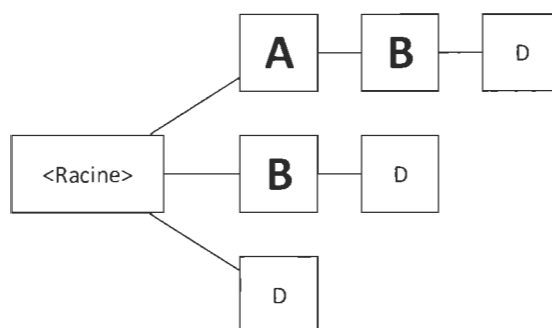
2.4.4.3 Comparaison/Soustraction

La comparaison de sujets se fait par la soustraction d'arbre A vers B, pour chaque nœud de A, on additionne la valeur de son compteur au maximum de poids de l'arbre et on vérifie son existence dans B, si le nœud de A existe dans B, nous attribuons des poids selon la méthode choisie d'attribution de ces poids. Deux exemples de soustraction/comparaison sont disponible à la Figure 22 et Figure 23 utilisant la méthode de poids uniforme. Cette comparaison est effectuée récursivement pour l'obtention du rapport entre le score obtenu et un score maximal. Au moment du calcul il est plus efficace de compter les poids perdus, soit les poids n'ayant pas de comparatif dans B, alors on obtient l'inverse du résultat souhaité à la fin. Soit un score de différence (Poids perdu sur poids maximum), un texte ayant moins de 30% de différence pourrait être considéré comme un même sujet, et un texte de moins de 10% peut être considérés très similaire, donc surement un extrait ou une citation.



Somme des pondération = 6

—

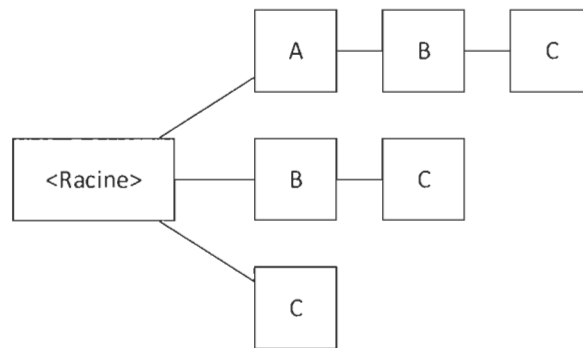


Somme des nœuds commun = 3

=

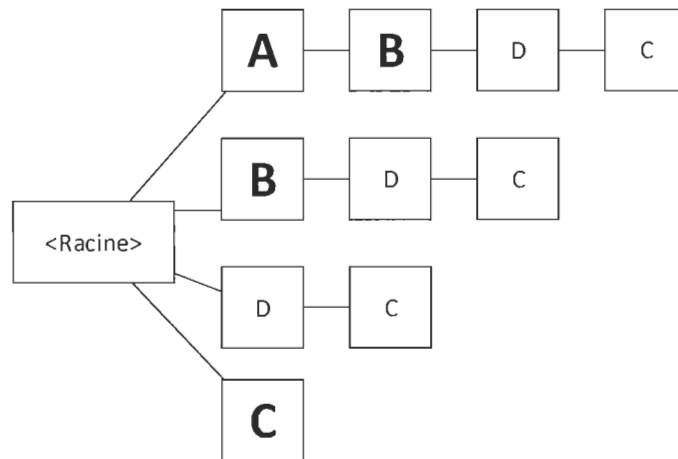
Similitude:
3/6

FIGURE 22 SOUSTRATIONS D'ARBRE



Somme des pondération = 6

—



Somme des nœuds commun = 4

=

Similitude:
4/6

FIGURE 23 SOUSTRATIONS D'ARBRE DE TAILLE DIFFÉRENTE

2.4.5 Amélioration de l'algorithme en arbre

Il est possible d'augmenter l'importance d'un mot en y substituant tous ses synonymes par lui-même. Par exemple remplacer « voiture » par « automobile » permettrait de reconnaître plus de répétitions des patrons identifiés avec automobile par exemple : « prendre sa voiture » se retrouverait jumelé avec « prendre son automobile » qui représentent en réalité la même expression.

Puisque N (Nombre de mots dans le texte) grandit très rapidement avec la longueur du texte et qu'il est possible d'analyser des paragraphes ou des textes entiers, nous pouvons tout en gardant la même approche réduire considérablement N en coupant le texte en phrase. À chaque phrase, nous éliminons toutes les têtes d'écriture. Cela permet de ne pas identifier une suite de mots chevauchant deux phrases distinctes, car cela ne représente pas un ajout d'information. Cependant, nous perdons aussi un peu de la notion de suite dans le texte. Il ne sera donc pas possible d'identifier « qu'un chasseur à tuer le lapin avant de le manger ». Cependant lorsque nous comparons des sujets, l'ordonnancement n'est pas ce qui est le plus intéressant. Nous perdons seulement le fait que le texte raconte une histoire, mais pas son sujet.

Dans le cas de réduction de la longueur avec les points de fin de phrase, le meilleur des cas devient plus évident : A. A. A. A. A. A. ... Cette suite permet de faire un arbre à un seul nœud « A » de N répétitions qui se parcourt en $O(N)$ plutôt que $O(N^2)$. La comparaison avec un autre arbre devient alors en $O(1)$. Une variante est de voir ceci comme un apprentissage séparé des arbres des phrases avec addition des arbres à la fin de cet apprentissage. Cependant la méthode par arbres distincts fera plus d'opérations dues à la création de nœuds qui existaient déjà dans le premier arbre et les opérations d'ajout ensuite. L'apprentissage direct dans l'arbre précédent est moins gourmand en ressources.

2.4.6 Différences majeures par rapport à l'algorithme des N-Gram

En ce qui a trait aux différences, elles sont positives ou négatives. Pour commencer, il sera question des différences négatives. Si et seulement si on utilise des N-Gram de taille fixe comparée à utiliser toutes les tailles possibles, l'algorithme en arbre peut alors devenir plus lourd que celui des N-Gram du fait qu'il aura encore le pire des cas à $O(N^3)$ tandis qu'une longueur fixe portera les N-Gram à $O(N^2)$. Cette différence pourrait laisser croire que l'algorithme en arbre serait plus long, mais pour une longueur de texte raisonnable, et étant donné que les comparaisons sont moins coûteuses dans l'arbre, ce temps pourrait être semblable. Maintenant, en ce qui a trait aux points positifs de l'algorithme en arbre comparativement aux N-Gram, une grande amélioration est la taille de l'arbre stocké en mémoire. En effet, puisque l'algorithme en arbre enregistre les mots comme nœud, il est alors plus léger à stocker que les N-Gram puisque dans ce contexte un nœud peut faire partie de plusieurs N-Gram. Le fait que l'arbre utilise uniquement un passage sur les données permet aussi de faire le traitement sans avoir à se rappeler de ce qu'on a déjà traité partiellement. Dans la présente recherche, en utilisant les pentes des courbes de suivi oculaire plutôt que les données brutes comme symbole de base et en permettant un degré de liberté (Voir Tableau 2 Choix CB) sur ceux-ci, il est alors aussi possible de détecter le même changement à différents niveaux dans un signal et être alors indépendant de la translation sur l'amplitude de ce même signal.

2.4.7 Généralisation de l'algorithme

L'algorithme en arbre a été élaboré afin de traiter les langages naturels mais peut être appliqué à une séquence d'entrée quelconque en adaptant quelques concepts. En gardant la notation de texte, il sera plus facile d'expliquer cet algorithme. Un mot sera alors l'unité de base de l'arbre, soit une courte série de données (lettres) ayant une signification. Comme dans le traitement des langues naturelles, il est possible de rouler l'algorithme directement sur les lettres, cela engendre un arbre beaucoup plus profond mais ayant une liberté plus grande quant aux fautes et synonymes. Une phrase est une série de mots se suivant dans la source de donnée. La source de donnée est alors appelée texte. Seule la notion de distance entre les mots est alors nécessaire à définir complètement.

2.5 Conclusion

À partir des informations recueillies dans plusieurs textes introduits dans la revue de littérature et une idée qui a émergé sur la possibilité d'utiliser un algorithme de détection des facultés affaiblies basée sur les N-Grams, il est maintenant possible de décrire avec plus de détails au chapitre 3, les méthodes expérimentales implémentées pour la détection des facultés affaiblies. Le chapitre 3 permettra dans un premier temps de rendre les tests symptomatiques basés sur la détection des réflexes oculaires comme le HGN, la dilatation/contraction et la convergence oculaire, possible dans un environnement virtuel mieux contrôlé et ainsi permettre d'évaluer les facultés d'une personne.

CHAPITRE 3

MÉTHODE EXPÉRIMENTALE

3.1 Introduction et hypothèses de base

Tout d'abord, le dispositif utilisé pour la capture et l'affichage des scènes dans l'environnement virtuel 3D pour les tests découle d'améliorations des dispositifs précédemment développés lors de recherches précédentes sur le même sujet. Un premier prototype a été conçu par Pierre Drolet [4]. Le dispositif utilisé dans cette recherche ne présente que la principale motivation du dispositif original qui était de faire la détection du HGN (nystagmus). Ainsi, la taille et le fonctionnement du nouveau dispositif n'ont pas d'égal dans le dispositif original. Les améliorations au dispositif de capture seront décrites en détail dans les prochaines sections. Le modèle de détection est abordé en premier du fait de sa complexité.

Pour en arriver au modèle d'identification des facultés affaiblies par un apprentissage automatique proposé, il a fallu émettre certaines hypothèses et tester le comportement d'un système avec ces différentes hypothèses. La première hypothèse nécessaire pour faire cette recherche est que n'importe quel signal se comporte comme un texte, soit une série d'évènements représentant similairement une série de mots. Cette hypothèse nous permet d'utiliser les algorithmes généralement utilisés en traitement des langues naturelles dans le domaine des facultés affaiblies. Ces algorithmes ont été choisis, parce qu'ils traitent des données qui changent dans le temps (position dans la phrase) et qui sont de tailles inconnues, un test ne peut donc pas toujours être résumé à des données sommaires.

La deuxième hypothèse est en fait une série d'hypothèses dont la validité devra être démontrée par la suite.

- Un mot est similaire à une pente dans un graphique (Identification des points d'inflexion). Pente qui pouvait dans le contexte de notre recherche représenter la variation de la direction de visée d'un suivi oculaire.
- Un mot a toujours une durée fixe dans le signal

Du fait même de ces hypothèses, il faut aussi spécifier qu'une même pente (identifié par les points d'inflexion ou la durée) à deux moments différents dans le texte aura la même signification primaire (voir Figure 24), la signification secondaire dépendra de son entourage.



FIGURE 24 EXEMPLES DE MOTS ÉGAUX, SIMILAIRES ET DIFFÉRENTS

3.2 Application de l'algorithme en arbre à des données de type autres que texte

L'algorithme en arbre identifie simplement une série de symboles(mots) utilisés dans un enregistrement (texte), pour l'appliquer à un enregistrement autre que du texte, par exemple le suivi du mouvement de l'œil que nous voulons effectuer dans cette présente recherche, il est possible de substituer la notion de mot par un mouvement, par exemple un mouvement de 1pixel/image vers le haut pendant une certaine durée, alors il suffirait de bien choisir la séparation entre ces symboles, soit une durée fixe, telle qu'un symbole extrait par 2 images (deux images sont nécessaires pour permettre le calcul de la pente) afin d'en extraire le plus de signification. Il est aussi possible d'utiliser directement la donnée brute comme symbole(lettre), il suffit de prendre la séparation qui donne les meilleurs résultats dans notre domaine. Les symboles peuvent être définis par des couples (Pente; Longueur) tels que (5; 10) ou (-1; 10).

Les améliorations possibles deviennent donc, d'identifier quelles pentes représentent la même information, similaire à la notion de synonyme. Puisque nous sommes maintenant en traitement de signaux plutôt qu'en traitement de langue naturelle, il n'est pas possible de connaître de manière absolue ce qui représente deux mêmes mots. La notion de mot n'est même pas claire. Ce qui rend plusieurs hypothèses nécessaires afin de pouvoir traiter notre information.

Dans le cas des déplacements de l'œil, nous désirons qu'une translation dans le signal n'affecte pas le résultat. Cela peut être réalisé en utilisant les pentes comme définition de mot au lieu d'utiliser les données brutes comme mot. En utilisant les pentes comme mot, il devient alors possible d'émettre plusieurs hypothèses sur la notion de synonyme et de distance. Ces hypothèses ont été énumérées dans la section 3.1 Introduction et hypothèses de base.

3.3 Méthode de tests des différentes solutions possibles

Dans cette section, il sera question de présenter la méthode d'analyse, soit le processus d'utilisation de l'appareil de réalité virtuelle et de capture des signaux, ainsi que son développement et une partie de son implémentation logicielle.

3.3.1 Méthode d'acquisition des données

Pour chaque type de substances que nous désirons pouvoir détecter, il faudra ajouter un test dit de connaissance à l'apprentissage. Les tests d'apprentissage sont les mêmes que les tests d'identification déjà existants à l'exception que l'on connaît déjà la catégorie.

3.3.1.1 Étapes d'un test

Pour ajouter un test à l'apprentissage, il suffit de suivre quelques étapes simples :

1. Ouvrir le programme de capture
2. Entrer le numéro de dossier selon le format
« TYPE_DE_SUBSTANCE » POINT « DOSSIER »
3. Effectuer les 3 tests (HGN, dilatation/contraction, convergence) pour l'œil gauche
4. Effectuer les 3 mêmes tests pour l'œil droit

Un dossier contiendra tous les tests, ceux contenant un point dans leur identifiant sont un apprentissage et ceux qui n'en ont pas seront ceux à identifier.

Certains prérequis sont cependant nécessaires pour réussir, d'abord, le casque (Oculus Rift) doit bien être positionné sur la tête pour empêcher les maux de cœur et permettre une détection idéale des yeux par rapport à la position des caméras. Nous devons avoir accès à une prise d'alimentation pour les caméras et d'un ordinateur capable d'utiliser un Oculus Rift « VR Ready ».

3.3.1.2 Le dispositif

Afin d'avoir un environnement contrôlé, il a été convenu d'utiliser un casque de réalité virtuelle. Il a d'abord été nécessaire de faire le pour et le contre de chaque casque disponible sur le marché. Notre premier choix était la lunette Microsoft HoloLens, ayant des écrans transparents, il aurait alors été possible de placer la caméra en face de l'utilisateur. Cependant ce casque n'était encore qu'en version alpha et non disponible au public. Les casques fonctionnant à l'aide d'un téléphone ont été omis, car un téléphone ne pourrait pas contenir vu sa mémoire réduite l'ensemble de la banque de connaissances. Le casque choisi a alors été l'Oculus rift puisqu'il s'intégrait directement dans le programme Unity que nous désirions utiliser pour le design des scènes 3D permettant la génération des stimulus visuels projetés aux yeux, ce qui nous a permis d'abstraire toute l'implémentation de ce monde 3D et de nous concentrer sur la capture et l'analyse des données. Des caméras ont été ajoutées dans l'oculus pour permettre la capture des mouvements des yeux. Étant donné le peu de place disponible, nous avons choisi des caméras

miniatures de sécurité qui capturent autant le spectre visible que le spectre infrarouge. Ces caméras doivent être placées sur le côté du casque à une distance très proche de l'œil, ce qui peut rendre la capture difficile. Dans le design initial, une carte d'acquisition des données biométriques devait permettre d'ajouter d'autres données telles que la température, les données de contraction musculaires (EMG) et les données de pulsation du cœur (ECG), cependant il n'a pas été possible de faire fonctionner cette carte avant la fin de la recherche. Il a alors été convenu d'utiliser uniquement les données de suivis oculaires des caméras. Des images du dispositif seront disponibles dans la section des améliorations au dispositif de capture.

3.3.2 Méthode d'identification des différentes solutions

Pour faciliter la compréhension de ce mémoire, certaines définitions doivent d'abord être comprises. Dans le cadre de cette recherche, une vidéo est l'unité de base, soit le résultat de la personne qui suit les instructions. Un test est l'ensemble des vidéos de la même personne devant suivre le même mouvement généré pour un stimulus visuel qui lui est généré par le générateur de scènes 3D projetées dans le casque. Un individu, dans le sens statistique du terme, est l'ensemble de tous les tests effectués par une personne à un temps donné. L'ensemble d'apprentissages quant à lui constitué d'arbres définis par l'ajout de plusieurs individus, on additionne l'ensemble des arbres représentant cet individu.

Pour une meilleure reconnaissance du pouvoir de décision de l'arbre, une méthode de validation croisée exhaustive (*leave-one-out cross-validation* [25]) sera utilisée. Chaque individu sera comparé à l'ensemble des autres données de l'apprentissage et un résultat du nombre de tests bien identifiés sera ensuite calculé. Cette validation sera appliquée à chaque mode possible.

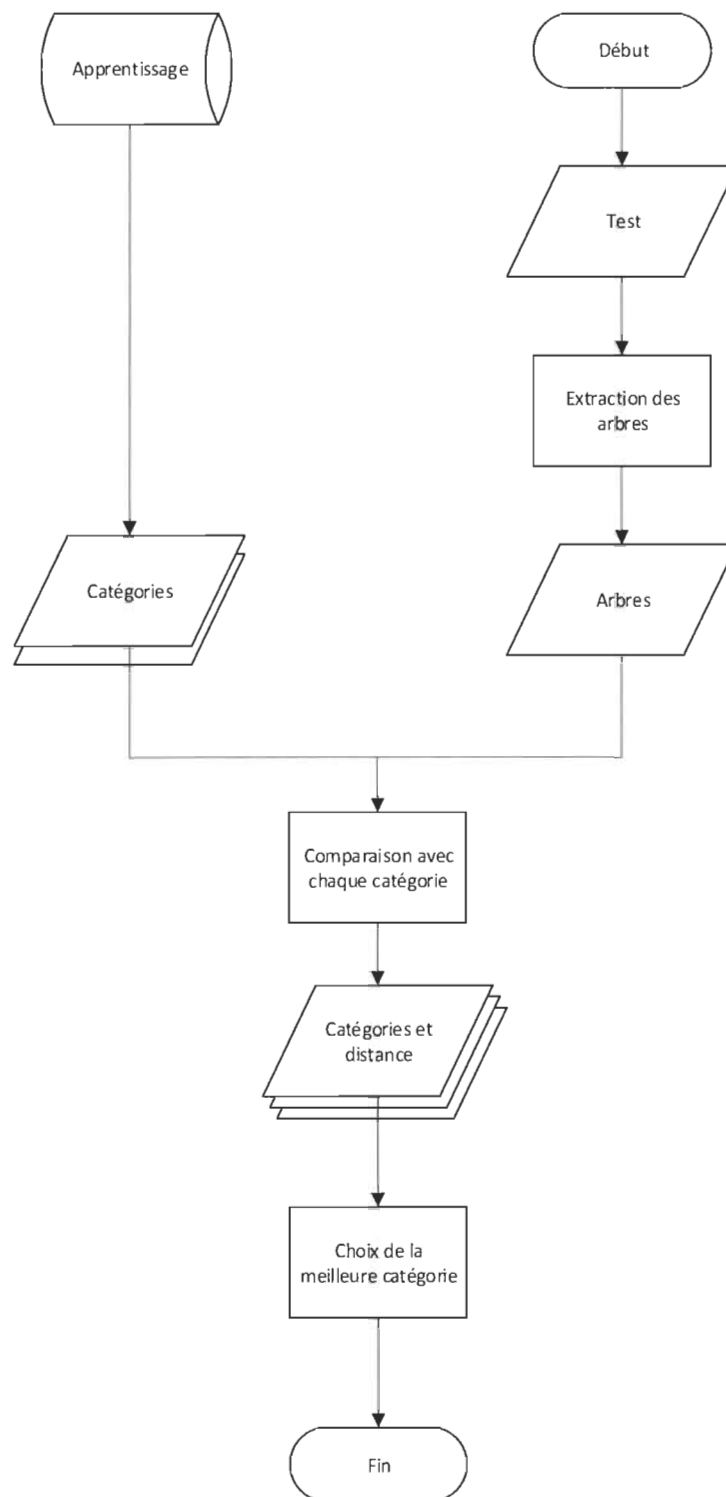


FIGURE 25 PROCESSUS D'IDENTIFICATION D'UN TEST

Dans le processus d'identification d'un texte (Figure 25), on débute avec les informations d'un test, soit l'ensemble des vidéos des deux yeux et des données biométriques si celles-ci ont été capturées. Un ensemble d'arbres extraits d'un test contient le même nombre d'arbres que le nombre d'un ensemble d'arbres représentant une catégorie de drogue (facultés affaiblies) dans l'apprentissage. Les profondeurs des arbres et leurs largeurs seront définies par les choix CA (Point de coupe) et CB (Degré de liberté) qui seront énumérés en détail plus tard. Une distance sera calculée entre chaque arbre de l'ensemble. Pour obtenir une mesure de distance entre les catégories de drogue, il faudra ensuite appliquer un processus d'unification de ces distances, ce processus sera défini par le choix CC (Unification des sources de données). Une fois les distances calculées, chaque test sera catégorisé selon la distance minimale. Cependant pour classer un individu, il faudra procéder au processus pour tous les tests disponibles, soit le test de nystagmus, de convergence et de dilatation/contraction. Un processus d'unification sera alors appliqué ici, aussi afin de prendre en compte les différents tests. Il sera défini par le choix CD (Unification des tests).

La définition de chaque choix sera détaillée dans la section 3.4.1 Différents calculs possibles.

Idéalement, pour empêcher que le test soit biaisé, certaines catégories de faux tests contenant un seul test devraient être ajoutées avec des données improbables. Les tests associés à ces catégories correspondraient alors à des personnes qui ne respectent pas les consignes ou une série de données aléatoire générée par ordinateur dans le but de vérifier le pouvoir d'identification de l'arbre. Ces tests uniques dans une catégorie non désirée servent à ce que le programme puisse répondre autre chose que les tests de base au cas où une personne ne serait pas en état de faire le test et donc qu'il ne serait pas concluant de répondre ni par « normal » ou « facultés affaiblies ».

3.4 Calcul et choix de la meilleure solution

Étant donné que la notion d'arbre s'applique à la base à des textes de langues courantes, l'appliquer à des données numériques sera un peu ardu. Certains choix devront donc être pris afin de faire un parallèle entre le monde numérique et textuel. Ces choix seront décrits dans cette section. De plus, chaque choix sera utilisé dans la recherche pour identifier celui qui a le meilleur résultat prédictif.

3.4.1 Différents calculs possibles

Pour donner suite au choix d'utiliser l'algorithme en arbre, il y avait encore des choix à effectuer, tels que choisir la notion de mot, de distance entre les mots, comment gérer des résultats de plusieurs sources et finalement comment identifier l'état de l'utilisateur avec ces différents tests. Quatre choix seront abordés et seront nommés CA, CB, CC et CD selon les lettres de l'alphabet et de leur ordre dans le processus.

Pour le premier choix, soit la définition d'un mot dans l'arbre, nous avons le choix de diviser les données résultant des tests en section selon les changements de pente des courbes de suivis oculaires, identifiées par la dérivée numérique ou de diviser le texte en section de longueur identique. Les deux options seront explorées plus tard dans ce document. La méthode par la dérivée est très sensible au bruit, mais a l'avantage de produire des arbres moins profonds dus aux mots qui seront généralement plus longs. La méthode de durée fixe équivaldrait à utiliser les lettres plutôt que les mots, elle peut être plus efficace en détection, mais est beaucoup plus lourde en calcul dû au nombre élevé de mots créés. Les deux méthodes ayant un bon potentiel, elles seront toutes deux expérimentées. Pour de meilleurs résultats, nous différencierons le choix de diviser en longueur fixe en 3 longueurs possibles pour un total de 4 choix. (Tableau 1)

1. Longueur variable (dérivée)
2. Longueur 2 images (Plus petite unité)
3. Longueur de 10 images (1/3 de seconde)
4. Longueur de 30 images (1 seconde)

TABLEAU 1 CHOIX CA (POINT DE COUPE)

Par exemple, pour un signal montrant une gaussienne, en longueur variable il pourrait il y avoir la série de couples (pente; durée) : [(0; 10) (5; 3) (-5; 3) (0; 10)] représentant une surface plane suivie d'une montée et d'une descente. Ce même signal peut être divisé selon des longueurs fixes pour obtenir par exemple [(0; 2) (0; 2) (0; 2) (0; 2) (0; 2) (5; 2) (0; 2) (-5; 2) (0; 2) (0; 2) (0; 2) (0; 2) (0; 2)] pour une longueur de deux images ou [(0; 10) (0; 10) (0; 6)] pour une longueur de dix images. Dans le cas de l'exemple d'une durée de 10 images, les valeurs à la position dix et à la position vingt ont la même longueur rendant la pente à 0 et ignorant totalement la gaussienne. Le dernier segment n'a pas la même longueur, car le signal s'arrête avant d'atteindre la fin de ce mot.

Pour le deuxième choix, la distance entre les mots, la formule de $d = 100d_s + d_L$, où d_s = différence entre les pentes et d_L représente la distance entre les longueurs de mots, sera prise en compte, peu importe la notion de mots. Dans le choix du mot de longueur fixe, on compte uniquement les pentes dans la distance puisque d_L sera toujours 0. Cette formule a été choisie, car nous partons de l'hypothèse que la pente est plus importante que la durée et que la hauteur du signal (Translation sur l'axe des données) n'a aucune importance. Encore une fois une étude plus approfondie des différentes méthodes de calcul de différence pourrait apporter un meilleur résultat de détection. Deux mots dont la distance est inférieure au seuil seront calculés comme égaux. Il sera question d'essayer différentes valeurs comme seuil de similarité, ces valeurs seront des puissances de 10 afin de parcourir une grande plage en un minimum d'itération. (Tableau 2)

1. 1
2. 10
3. 100
4. 1000

TABLEAU 2 CHOIX CB (DEGRÉ DE LIBERTÉ)

Comme troisième choix, soit comment faire la gestion de plusieurs sources de données, plusieurs choix s'offrent à nous. Voici les choix qui seront abordés dans ce mémoire (Tableau 3)

1. Moyenne des différents résultats
2. Le minimum des différents résultats
3. Le maximum des différents résultats
4. La multiplication de tous les résultats
5. Moyenne pondérée des différents résultats

TABLEAU 3 CHOIX CC (UNIFICATION DES DIFFÉRENTES SOURCES DE DONNÉES)

Chaque type de calcul a ses avantages qui seront maintenant explorés. En premier lieu, la moyenne est le moyen le plus simple de gérer plusieurs résultats, elle n'est pas très affectée dans le cas de résultats aberrants. En deuxième lieu, le choix du minimum permettrait de dire que les différentes sources de données sont toutes discriminantes, soit dès qu'une donnée est différente alors le sujet est différent. Au contraire, le maximum permet de dire que les données sont discriminantes de manière positive, soit que chacune a un pouvoir d'identification et que le résultat le plus similaire représente toujours la catégorie de drogue peu importe le résultat des autres données. La multiplication des valeurs peut aussi représenter la discrimination du fait qu'une similitude proche de 0 amènera le résultat proche de 0. Cependant, cette méthode donnera quasiment toujours des résultats très bas. Il suffit de penser à une valeur entre 0 et 1 à la puissance du nombre de données, cela diminue la valeur considérablement. La dernière option est une moyenne pondérée, elle pourrait de son côté représenter que la mesure de pupille est plus importante dans le test de dilatation que la mesure d'angle.

Pour différents résultats, nous prenons en compte plusieurs facteurs, pour faciliter le traitement, sachant de plus qu'un arbre considère les données indépendamment. Il y a 10 sources de données possibles présentées dans le Tableau 4.

Donnée	Plage possible	Description
X	0-640	Position en X (centre) de la pupille dans l'image
Y	0-480	Position en Y (centre) de la pupille dans l'image
Largeur	0-640	Largeur de la pupille en pixel
Hauteur	0-480	Hauteur de la pupille en pixel
Diamètre MM		Diamètre de la pupille en millimètre, calculé selon la taille de l'iris identifié.
Angle	-45 - +45	<p>Angle calculé selon la forme de la pupille, selon le positionnement de la caméra, une pupille ronde = 45°* soit directement vers la caméra, une pupille qui regarde vers l'avant 0° serait 2 fois plus haute que large et une pupille qui regarde du côté opposé serait une fine ligne (Proche de 0 de largeur)</p> <p>*45° représente le côté droit de la personne, si on utilise la caméra gauche, regarder la caméra nous donnera - 45° degrés plutôt.</p>

Largeur Iris	1-640	Largeur de l'iris en pixel, doit être supérieur à la largeur de la pupille
Hauteur Iris	1-480	Hauteur de l'iris en pixel, doit être supérieure à la hauteur de la pupille
Pupille trouvée	0-1	0 indique que la pupille n'a pas été trouvée sur cette image et 1 la présence d'une pupille
Iris trouvé	0-1	0 indique que l'iris n'a pas été trouvé sur cette image et 1 la présence d'un iris

TABLEAU 4 LES DIFFÉRENTES DONNÉES PROVENANT D'UN VIDÉO

Les données brutes telles que **X, Y, Largeur et Hauteur** dépendent de la résolution de l'image, dans l'exemple, nous prenons en compte une image de 640x480.

Un dernier choix s'offre à nous, puisque ce dernier choix nous donne un résultat par test (Nystagmus, Convergence et Dilatation/contraction), il faut jumeler les résultats finaux. Plusieurs méthodes s'offrent à nous. Tout comme le choix précédent, il sera question d'appliquer le minimum, le maximum, la multiplication et la moyenne sur les résultats de similitude. Cependant, il y aura aussi d'autres choix possibles tels que : devons-nous accepter que si 2 tests pris de manière individuelle donnent le même résultat, alors l'usager est dans cet état? Nous pouvons aussi appliquer cette même logique à 3 tests sur 3 qui donnent le même résultat. Un autre choix s'offrant à nous est de calculer la similitude avec un état en prenant sa position dans les résultats, la première position vaudrait un point, ensuite un demi-point, un tiers et ainsi de suite.

Test 1	Test 2	Test 3
Normal (1 pt)	Alcool (1 pt)	Alcool (1 pt)
Alcool (1/2 pt)	Normal (1/2 pt)	Normal (1/2 pt)
Cannabis (1/3 pt)	Cannabis (1/3 pt)	Cannabis (1/3 pt)
... (1/4 pt)	... (1/4 pt)	... (1/4 pt)

TABLEAU 5 EXEMPLE DE POINTS SELON LA POSITION

Dans cet exemple (Tableau 5), la catégorie « Normal » a obtenu 2 points, « Alcool » a quant à lui obtenu 2 points et demi, ce qui le positionne comme catégorie la plus plausible.

Pour résumer les méthodes possibles, voici la liste qui sera utilisée dans l'application CD=1 à CD=12 (voir Tableau 6)

1. Moyenne
2. Minimum
3. Maximum
4. Multiplication
5. Moyenne avec 2 choix activés
6. Minimum avec 2 choix activés
7. Maximum avec 2 choix activés
8. Multiplication avec 2 choix activés
9. Moyenne avec 3 choix activés
10. Minimum avec 3 choix activés
11. Maximum avec 3 choix activés
12. Multiplication avec 3 choix activés

TABLEAU 6 CHOIX CD (UNIFICATION DES TESTS)

3.5 Améliorations au dispositif de capture

Dans cette section, il sera question des améliorations apportées au dispositif de capture depuis le dernier prototype développé lors des dernières recherches, bien que certaines de ces améliorations aient été effectuées un peu avant le début de ce présent projet de recherche, s'agissant du même projet ces améliorations sont quand même incluses.

Il est bien de donner une idée des prototypes développés lors des recherches précédentes avant de pouvoir énumérer les améliorations faites dans la présente recherche. Dans les études précédentes, le dispositif était composé d'un écran permettant l'affichage de symboles blancs derrière une visière d'un casque de soudure avec un éclairage DEL. Ce dispositif n'était donc pas très réaliste pour l'utilisateur. La caméra était alors visible derrière la visière et elle captait le reflet de l'écran sur celle-ci. Le dispositif était donc encombrant pour l'utilisateur.

En premier lieu, il fallait donc, rendre le dispositif plus portable, soit en diminuant sa taille et les besoins en câbles pour connecter les différentes composantes électroniques, dont le système d'illumination DEL. Plusieurs options ont été envisagées telles que des écrans transparents pour éliminer la réflexion en permettant de placer une caméra à l'arrière de cet écran transparent, mais ces écrans sont difficiles à obtenir sur le marché et aussi difficiles à intégrer au système. Une autre option était d'utiliser un casque de réalité virtuelle comme l'Oculus Rift (Figure 26). Plusieurs types de ce genre de casques existent dont certains qui utilisent un téléphone comme écran. L'utilisation d'un casque de réalité virtuelle branché à un ordinateur a été pensée en conséquence d'être sur la tête d'un utilisateur sans être encombrante. Nous obtenions alors directement un dispositif plus portable et qui prend moins de câbles entre autres pour le branchement du système d'illumination, puisque dans le casque de réalité virtuelle l'illumination peut être contrôlée par l'ajustement de la luminescence de l'écran du casque. Ce casque a été choisi plutôt que celui avec un téléphone mobile pour mieux synchroniser le vidéo avec la détection.



FIGURE 26 CASQUE OCULUS

Il fallait ensuite rendre le test plus réaliste, en créant une scène 3D permettant de produire les stimulus visuels requis dans le suivi oculaire des tests HGN, dilatation/contraction et de convergence. Il était alors possible pour l'utilisateur d'avoir l'impression que les tests sont « réels » (3D) plutôt que de suivre une ligne (2D) blanche. Par « réel » nous voulons dire plus proche des tests DRE ou d'un test pouvant se passer dans le vrai monde. Suivre un cube que l'on voit en trois dimensions est plus réaliste que suivre un carré coloré sur un écran 2D. Cependant il serait possible avec ce type de casque de faire une scène 3D complète comme le modèle d'un policier sur le bord de la route tenant un dispositif de suivi pour le test. Cette option fut celle choisie,

cependant il valait mieux obtenir des résultats que d'obtenir de belles scènes pour un prototype. Ces scènes peuvent facilement être créées par la suite étant donné que la scène n'affecte pas le test à l'exception du test de dilatation dû à l'illumination.

Une fois le monde virtuel construit, il a fallu ajouter des fonctionnalités de capture de données, sans quoi aucun test n'aurait pu être réalisé. Étant donné l'environnement compact d'un Oculus Rift, il a fallu concevoir des socles sur mesure (voir Figure 27) pour accueillir les caméras. Ces socles ont alors été imprimés en 3D avec une imprimante 3D de la bibliothèque de l'UQTR. Aussi dû à leur taille, les caméras ne pouvaient pas être positionnées devant les yeux, seulement sur le côté du casque, ce qui rend l'œil non visible lorsqu'il regarde dans la direction opposée. Il

est cependant possible de calculer l'angle d'une pupille absente de champs de vision de la caméra en utilisant l'angle de la pupille opposée. En ce qui a trait à l'illumination, les caméras sont munies d'un cercle d'illumination NIR. (Voir Figure 26 Casque Oculus et Figure 28)

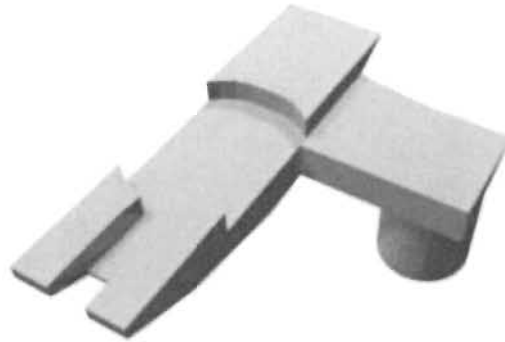


FIGURE 27 SOCLES POUR CAMÉRAS AVANT IMPRESSION

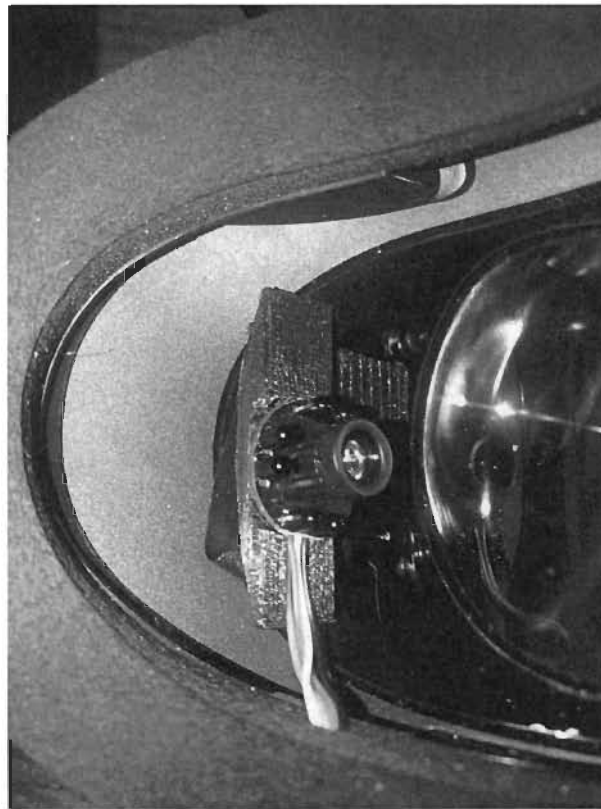


FIGURE 28 DISPOSITIF DE CAPTURE DE SÉQUENCES VIDÉO PAR CAMÉRA NIR

Une fois le dispositif matériel conçu, il fallait extraire les données à partir des vidéos. Pour ce faire, un processus en plusieurs étapes a été utilisé. Il est quelque peu similaire au processus d'identification de l'œil développé dans les recherches antérieures [6], mais a été simplifié. La capture des images à partir de la caméra ou d'un vidéo enregistré se fait à l'aide de EMGU.CV [13], une librairie de traitement d'image.

La première étape consiste à uniquement obtenir le vidéo et de l'enregistrer sur le disque. Une fois le test complètement enregistré, l'analyse commence sur ce vidéo. Pour le traitement du vidéo (Figure 29), une fois la caméra démarrée, un certain nombre d'images sont rejetées pour permettre à obturateur des caméras de s'adapter à l'illumination ambiante. Ensuite, un algorithme K-Means [26] est appliqué ce qui réduit le nombre de couleurs de l'image à un nombre défini dans un fichier de paramètres et crée une palette de couleurs désirées. Dépendamment du type de caméra, ce nombre de couleurs peut varier. Pour chaque image suivante, cette palette de couleurs sera appliquée. Une sélection de couleur sera effectuée sur cette palette de couleur servant aussi d'histogramme. Uniquement un certain nombre de couleurs les plus foncées parmi cette liste seront utilisées pour la détection de la pupille, sachant qu'elle apparaît foncée dans les images. Il se peut que plusieurs zones correspondent au critère de seuillage. La zone avec la plus grande taille avec une proportion raisonnable et centrée sur la zone où devrait se trouver l'œil est sélectionnée.

La décision de ne pas prendre en compte les données extraites de l'image précédente a été prise à la suite de quelques tests. Ce choix a été fait, car dans bien des cas les premières images ne sont pas d'aussi bonne qualité, ce qui entraînait un mauvais positionnement de l'œil ce qui rendait ensuite l'identification erronée. Dans plusieurs tests, conserver l'information précédente menait à une mauvaise détection, car une erreur pouvait se répéter d'une image à l'autre.

L'iris sera ensuite identifié (Figure 30) autour de cette pupille par agrandissement successif par croissance de région jusqu'à détecter le blanc de l'œil par la moyenne du niveau de gris qui deviendra alors plus claire. Ce processus se répète jusqu'à à fin du vidéo. Le traitement énuméré se fait dans la librairie ImageProcessing [14] de source ouverte. Cette librairie permet

des actions similaires à OpenCV/EMGU.CV, mais en ajoutant des critères plus sensés tel que lors de la détection de contour en identifiant un critère de sélection plutôt que de laisser le critère à la discrétion de la librairie.

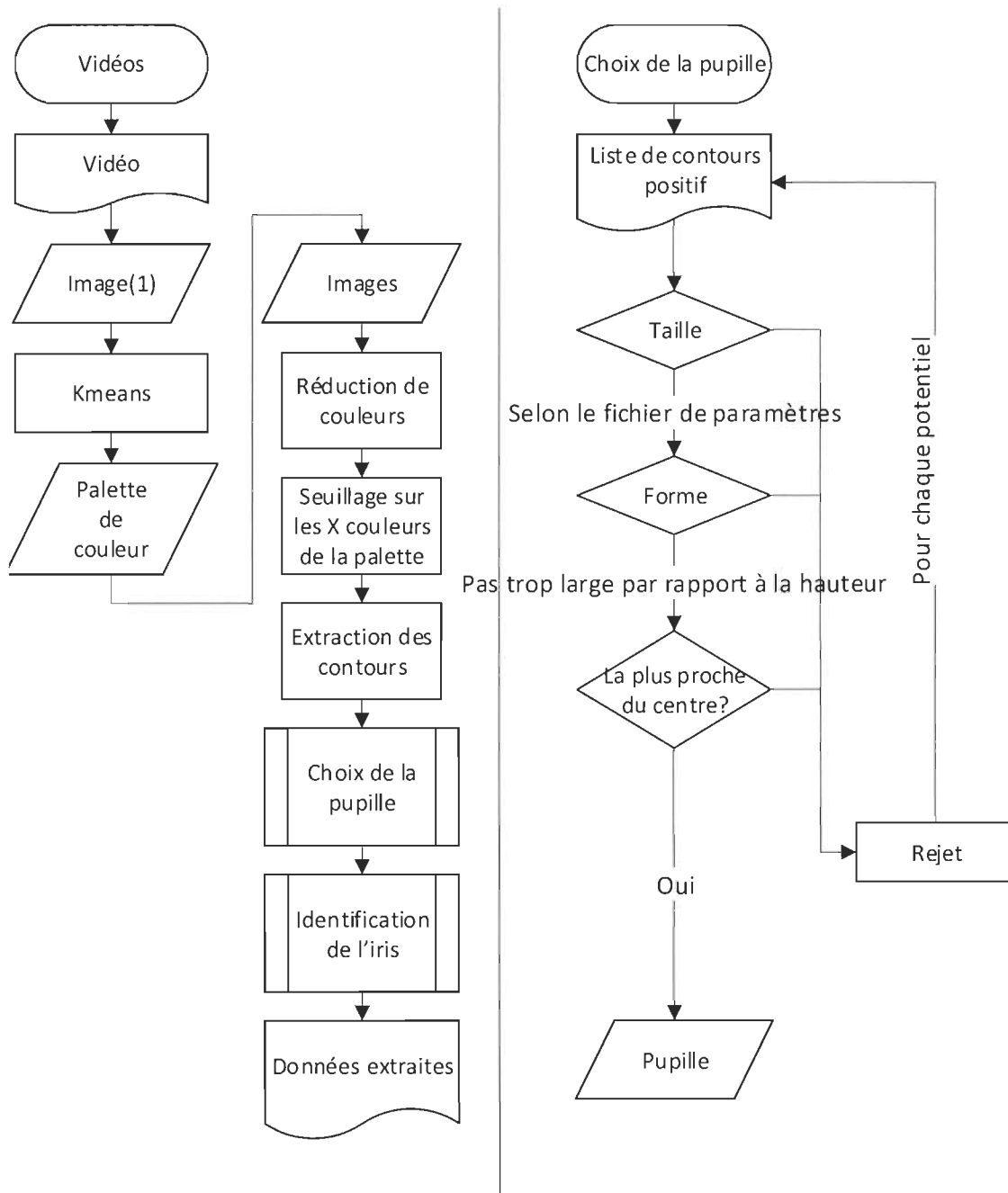


FIGURE 29 PROCESSUS D'IDENTIFICATION DE LA PUPILLE DANS LES VIDÉOS

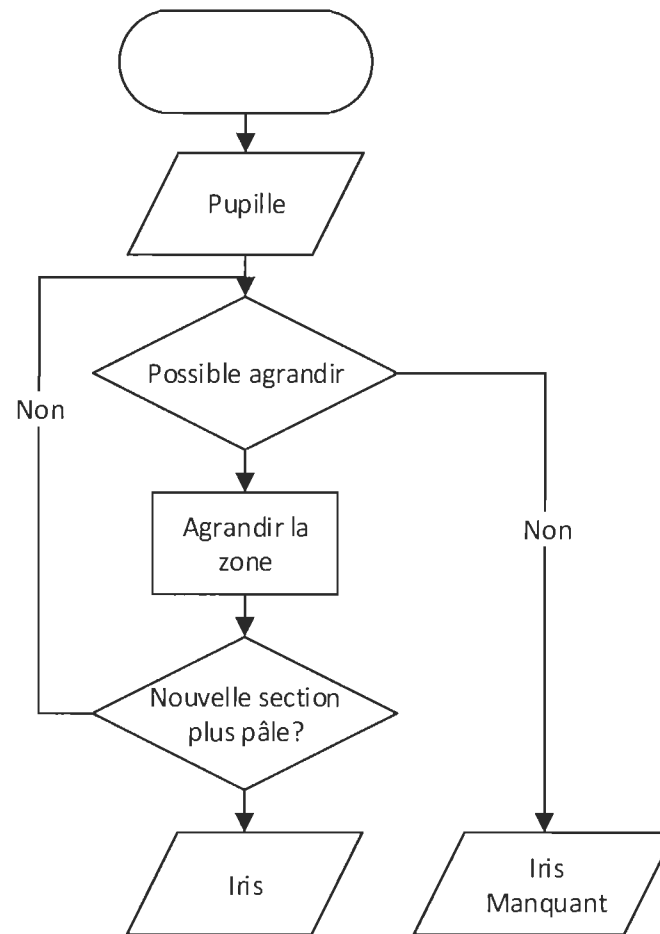


FIGURE 30 PROCESSUS EXTRACTION DE L'IRIS

Une fois les données de base extraites, vient le calcul des informations utiles à l'évaluation des facultés d'une personne. Ces informations incluent l'angle de visée de l'œil, et les mesures de taille de l'iris. L'angle de visée est calculé par le ratio hauteur sur largeur de la pupille. La formule suivante a été utilisée dans un premier temps. Si la pupille est bien ronde, l'œil regarde dans la direction de la caméra et plus l'œil devient une fine ligne, plus il regarde à l'opposé de la caméra.

$$\theta = 90 \frac{\text{Largeur}}{\text{Hauteur}} - 45$$

ÉQUATION 1 ANGLE DE VISÉE SELON RATIO

Après l'analyse de multiples vidéos, il a été convenu d'utiliser une polynômiale (Figure 31) de degré deux pour évaluer l'angle de visée pour donner suite à des expérimentations. Cette nouvelle formulation donne sensiblement les mêmes résultats, mais est plus précise puisque tenant compte de la courbure de l'œil plutôt que de faire une interpolation linéaire entre les extrêmes. Le rapport Largeur/Hauteur est restreint dans l'intervalle 0,1 puisque l'œil ne peut être plus large que haut sauf si la caméra induit une déformation.

$$\theta = -85.714 \left(\frac{Largeur}{Hauteur} \right)^2 + 175.71 \frac{Largeur}{Hauteur} - 45$$

ÉQUATION 2 FORMULE D'IDENTIFICATION DE L'ANGLE DE VISÉE CORRIGÉE

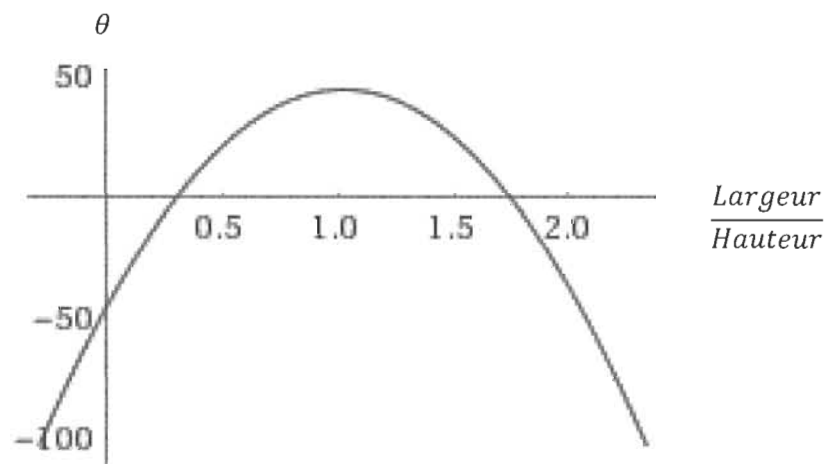


FIGURE 31 FORMULE D'IDENTIFICATION DE L'ANGLE DE VISÉE CORRIGÉE

*Utilisable sur]0,1]

Pour l'iris, la détection ne se fait pas par seuillage, mais plutôt par des agrandissements successifs (Croissance de région). Tout en gardant le même ratio que la pupille, on agrandit l'ellipse d'approximation de la pupille en gardant le même centre. Ensuite, on calcule le taux de gris moyen, si celui-ci est très supérieur à la valeur précédente c'est qu'on a atteint le blanc de l'œil. Quand l'iris a été détecté, la mesure du diamètre de la pupille est calculée à partir du diamètre de l'iris puisque l'iris est de taille connue (10mm). Si l'iris n'a pas été trouvé, la mesure de diamètre de la pupille doit être calculée selon un taux de conversion entre les millimètres et les pixels.

Pour de meilleurs résultats d'identification des facultés d'une personne, il a été décidé d'ajouter d'autres facteurs physiologiques humains tels que les battements cardiaques, les données de contractions musculaires et de température. Pour ajouter ces informations, il fallait concevoir une carte d'acquisition des données. Pour concevoir une carte, il fallait d'abord définir nos besoins. La carte devait pouvoir communiquer avec l'ordinateur en temps réel pour l'enregistrement des données et pouvoir capter ces différentes données. Comme premier choix pour cette carte, il a été convenu d'utiliser le composant "Onion Omega2", il permettait en même temps de connecter la carte d'acquisition au wifi, et faire le traitement des autres capteurs dû à sa puissance de calcul. En ce qui a trait à la capture des données musculaires, un capteur déjà conçu spécialement à cette fin a été retenu. Ce capteur possède déjà la configuration nécessaire pour obtenir les informations musculaires à partir de sondes placées sur la peau. La sortie de ce capteur est analogique, ce qui veut dire des données dont la tension électrique varie entre 0 et 3.3V. (Figure 32)



FIGURE 32 CAPTEUR DE DONNÉES MUSCULAIRE (EMG)

Étant donné que l'Omega n'a pas d'entrée analogique, il a fallu choisir un convertisseur A/D. Le choix de ce convertisseur a été orienté par plusieurs facteurs. Le premier facteur est le protocole de transport de données, dans le cas du SI8902B [27], la communication utilise le protocole SPI. Ce protocole est facile d'utilisation puisqu'un seul maître lance la communication, il gère la vitesse et communique toujours bi-directionnellement. Il utilise aussi uniquement 4 connecteurs pour communiquer. Ce convertisseur a aussi été utilisé dans un projet précédent ce qui éliminait les problèmes de fonctionnement.

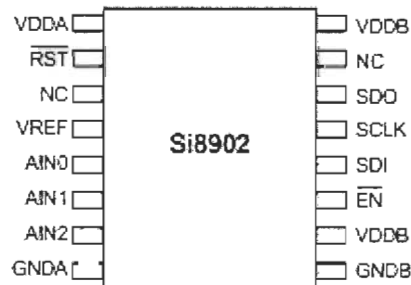


FIGURE 33 SI8902B RÉCUPÉRÉE DE SILICON LABS [27]

Sur la Figure 33 nous pouvons apercevoir le nombre d'entrées analogiques (AIN) et les entrées de communications SPI (SDO, SCLK, SDI, EN). Le reste des entrées permettent l'alimentation du circuit.

Le capteur cardiaque (Figure 34) a aussi été choisi pour sa simplicité d'utilisation. Il se place sur le bout d'un doigt afin de mesurer les battements cardiaques à l'aide d'une diode. En ce qui concerne la communication, celle-ci est effectuée aussi de manière analogique. Ce capteur est originalement conçu pour fonctionner avec des appareils de type Arduino, ce qui rend son utilisation très facile. Il suffit de l'alimenter pour qu'il commence à émettre des données.

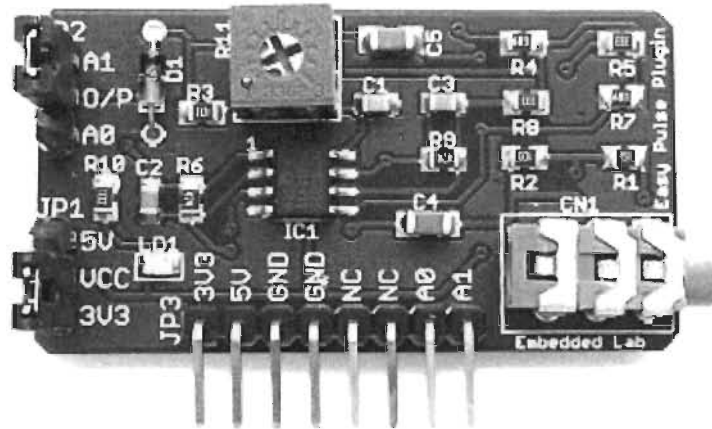


FIGURE 34 CAPTEUR CARDIAQUE [28]

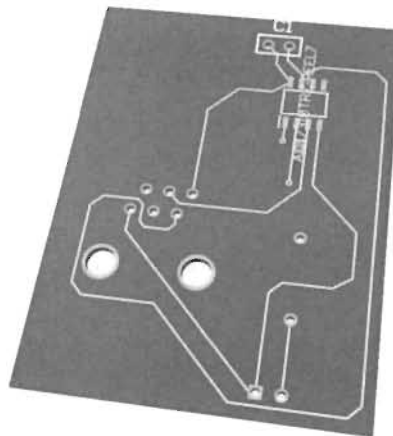


FIGURE 35 CARTE POUR ACCUEILLIR CAPTEUR DE TEMPÉRATURE

Le capteur de température a lui aussi été choisi pour sa taille et son moyen de communication. Il communique lui aussi selon le protocole SPI. Un capteur non encombrant devait être choisi pour ne pas altérer la réponse du sujet. N'ayant pas de capteur de température humaine non encombrant, un capteur de température ambiante a été choisi. Un capteur ambiant collé sur la peau humaine devrait donner la température de la peau, il est alors possible de calculer la température corporelle à partir de cette température de peau ou de baser notre base de connaissances sur les températures de peau. Ce capteur n'était pas très connu dans le

domaine, ce qui a rendu son utilisation difficile. Dans le projet, il était prévu d'utiliser ce dispositif de température de manière brute, ce qui veut dire d'obtenir les informations directement sans traitement pour les amener en Celsius et ainsi seulement faire apprendre ces données brutes au système. Une carte (PCB) a été conçue pour l'accueillir, et est visible à la Figure 35.

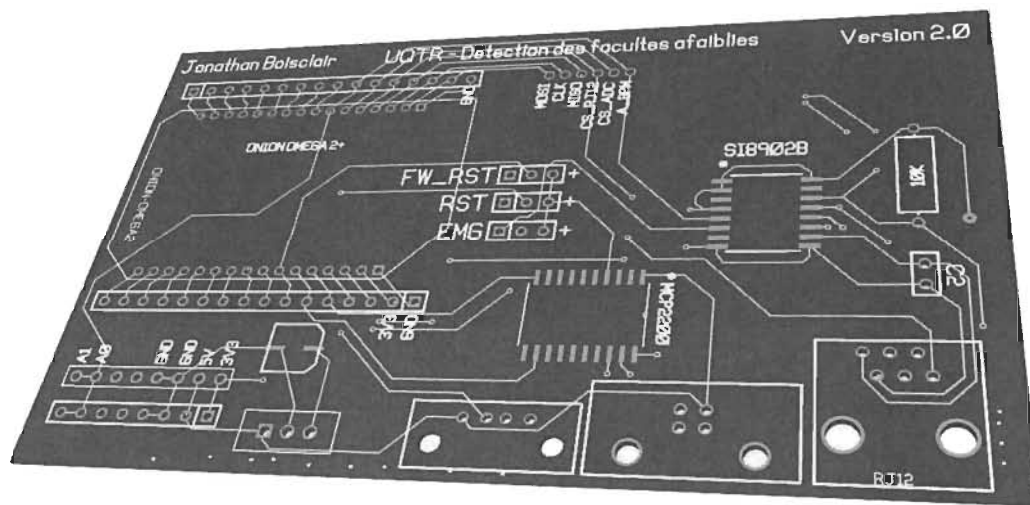


FIGURE 36 PCB ACQUISITION

Une fois la carte d'acquisition complètement assemblée (Figure 36), nous avons rencontré plusieurs problèmes, premièrement le capteur de température ne répondait à aucune demande bien que nous suivions le protocole de communication à la lettre. Ensuite, les données de cœur et de muscle étaient tellement bruitées, d'une cause inconnue, qu'il était impossible d'en sortir une seule information utile. Cette amélioration du dispositif a alors été mise de côté afin de prioriser les tests basés sur les réflexes oculaires utilisant les caméras. Il est possible que les bruits aient été causés par une alimentation défectueuse de la carte ou par une erreur de conception due au manque d'expérience.

3.6 Conclusion

Ayant précédemment identifié le processus par lequel les données seront extraites à partir des vidéos de suivi oculaire, il est temps d'exécuter ce processus et ainsi d'obtenir beaucoup de données permettant d'entraîner un classificateur qui permettra la détection des facultés affaiblies. Dans le prochain chapitre, nous expliquerons les données obtenues par les vidéos et par la suite les données provenant du traitement de détection de la substance consommée. Dû à un grand nombre de possibilités de traitements à la détection, nous devons en faire un résumé afin de les inclure dans ce mémoire.

CHAPITRE 4

RÉSULTATS, DISCUSSION ET INTERPRÉTATION

Dans les chapitres précédents, les méthodes de détection des yeux dans les vidéos de suivi oculaire et de construction des arbres de représentation de la connaissance ont été introduites. Dans ce chapitre, nous présenterons d'abord les résultats de détection des yeux et ensuite d'identification de la classe de symptômes associés à l'état des facultés d'une personne. Les résultats de la validation croisée seront présentés en premier et ensuite un résumé des décisions et conclusions présent par rapport à ces résultats.

Comme spécifié dans le chapitre précédent, les données des vidéos seront traitées par ordinateur afin d'en extraire l'information. Dû à la faible quantité d'information ayant pu être recueillie avec le dispositif ancré à une station non déplaçable, il n'a pas été possible de retirer les vidéos dont la qualité était inférieure. Ces vidéos de faible qualité ne pouvaient être traités par l'algorithme puisque par exemple, des cils et autres parties du corps cachaient l'œil. Il a donc été convenu que pour traiter ces informations, un humain allait donc extraire les informations afin de pouvoir exécuter l'algorithme en arbre sur une plus grande quantité de vidéos.

Bien que cette quantité de vidéos soit très basse pour pouvoir en extraire des conclusions définitives, une tendance pourra en être extraite afin d'obtenir une preuve de concept pour tester ce même algorithme sur une plus grande quantité de données. Pour ce faire des modifications devront être faites au programme pour stocker les arbres sur disque plutôt qu'en mémoire dû à la quantité plus imposante de données. Une limite pourrait aussi être imposée à la profondeur de l'arbre afin de calculer avec une quantité de données plus imposante et ainsi pouvoir trouver un point d'équilibre entre la vitesse, la quantité de données et la qualité des résultats obtenus.

4.1 Extraction des données à partir de vidéos

Les données extraites des vidéos seront présentées ici et ce découlant d'un test (Convergence) effectué sur une personne dont les facultés sont normales. Depuis, ces courbes (voir Figure 37) sont décrites pour en faire ressortir les caractéristiques qui sont extraites et analysées par l'algorithme de détection des facultés affaiblies.

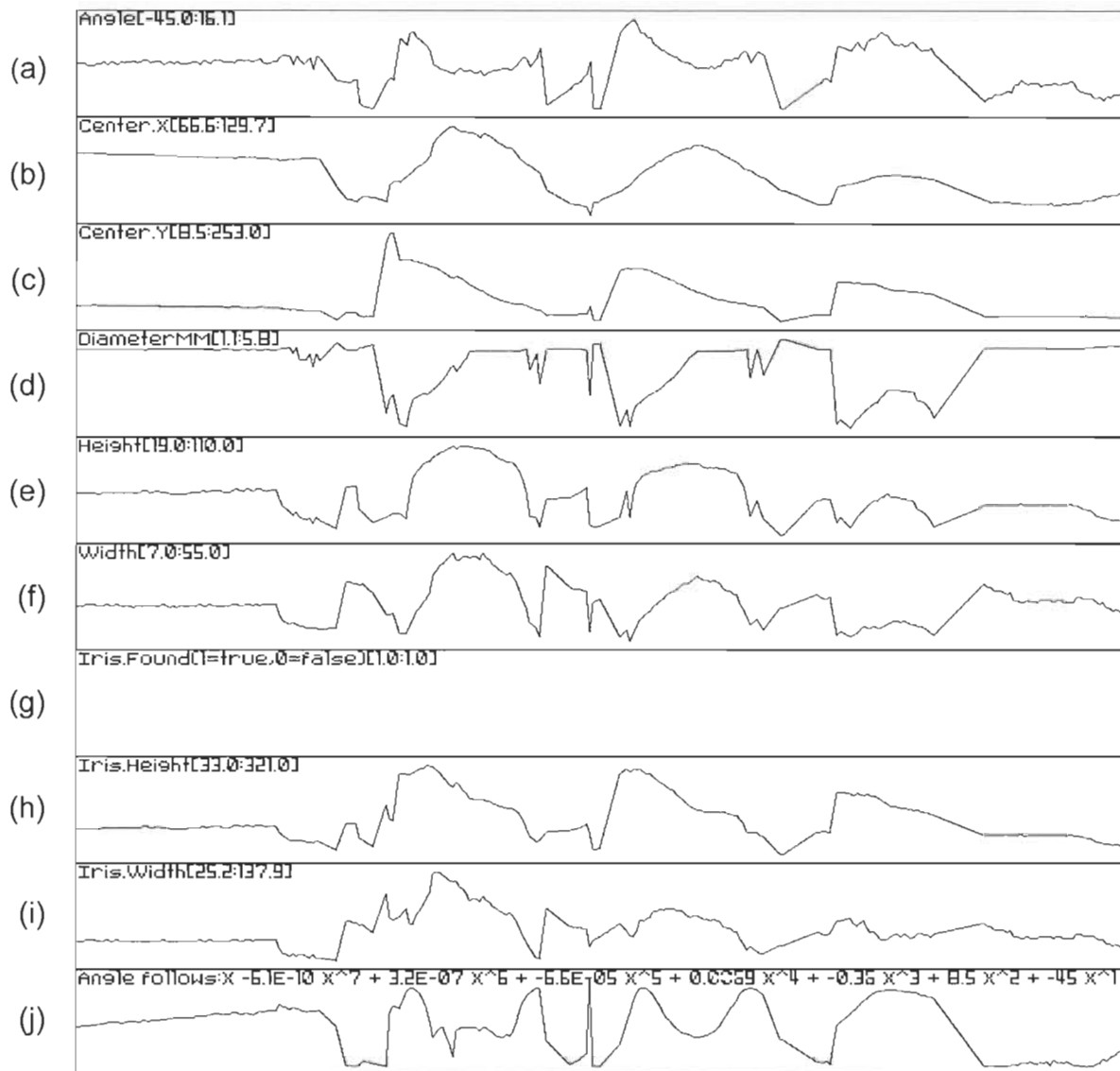


FIGURE 37 RÉSULTATS D'UN TEST DE CONVERGENCE DE L'ŒIL GAUCHE

(a) Angle de visée (b) Position x de la pupille (c) Position y de la pupille (d) Diamètre de la pupille (e) Hauteur en pixel de la pupille (f) Largeur en pixel de la pupille (g) Si l'iris a été détecté dans cette image (h) Hauteur de l'iris détecté (i) Largeur de l'iris (j) Angle de visée en utilisant une polynomiale pour l'adoucir.

Des résultats du test de convergence de l'œil gauche (voir Figure 37), il peut être conclu que le centre en X et en Y suivent la balle qui effectue un mouvement en spirale vers le centre dans l'environnement 3D du casque RV. Dans ce graphique (voir Figure 37 (c)), il est possible d'observer que le centre au repos n'est pas centré verticalement sur la courbe de mouvement. L'angle (voir Figure 37 (a)) aurait suivi un mouvement entre -45 et 16 degrés, ce qui n'est pas réellement le cas puisque la balle ne dépasse pas 30 degrés par rapport à la position de l'œil. De plus la mesure d'angle est beaucoup trop bruitée pour pouvoir en sortir une conclusion plausible, le choix de ne pas l'inclure dans l'arbre a été pris à la suite d'une très mauvaise détection lors de l'inclusion de celui-ci. Une légère dilatation a aussi été identifiée qui semblait avoir une corrélation avec le centre (voir Figure 37 (c, d)). Il est possible d'expliquer cette dilatation par le fait que l'objet utilisé pour générer le stimulus visuel était d'un rose assez vif dans une scène noire. Lorsque l'objet se trouvait dans l'axe optique de la caméra, l'œil se trouvait beaucoup plus éclairé que lorsque l'objet se trouvait à l'opposé caché par le nez. Il est possible aussi d'en déduire que l'iris a toujours été détecté (voir Figure 37 (g)). La Figure 37 (j) du graphique est supposée être l'angle lissé par une polynomiale de degré 7, cependant cet angle lissé ne donne pas des mesures plus précises que celui non lissé et il sera donc lui aussi exclu de l'apprentissage.

Il est possible d'extraire plusieurs informations du test de dilatation/contraction dont les résultats sont exposés à la Figure 38. Premièrement la personne regarde principalement toujours en avant puisque l'angle se situe entre -6 et 5 degrés (voir Figure 38 (a)). Le centre (voir Figure 38 (b, c)) change très peu correspondant alors à aucun changement de l'orientation de son regard. Il y a deux dilatations de la pupille (Voir Figure 38 (d)) qui peuvent être identifiées dans la durée du test qui est de 8 secondes avec 2 changements d'illumination. La pupille se contracte aussi très rapidement, mais n'est identifiée qu'une seule fois, car l'enregistrement se termine avant la deuxième contraction (voir Figure 38 (d)).

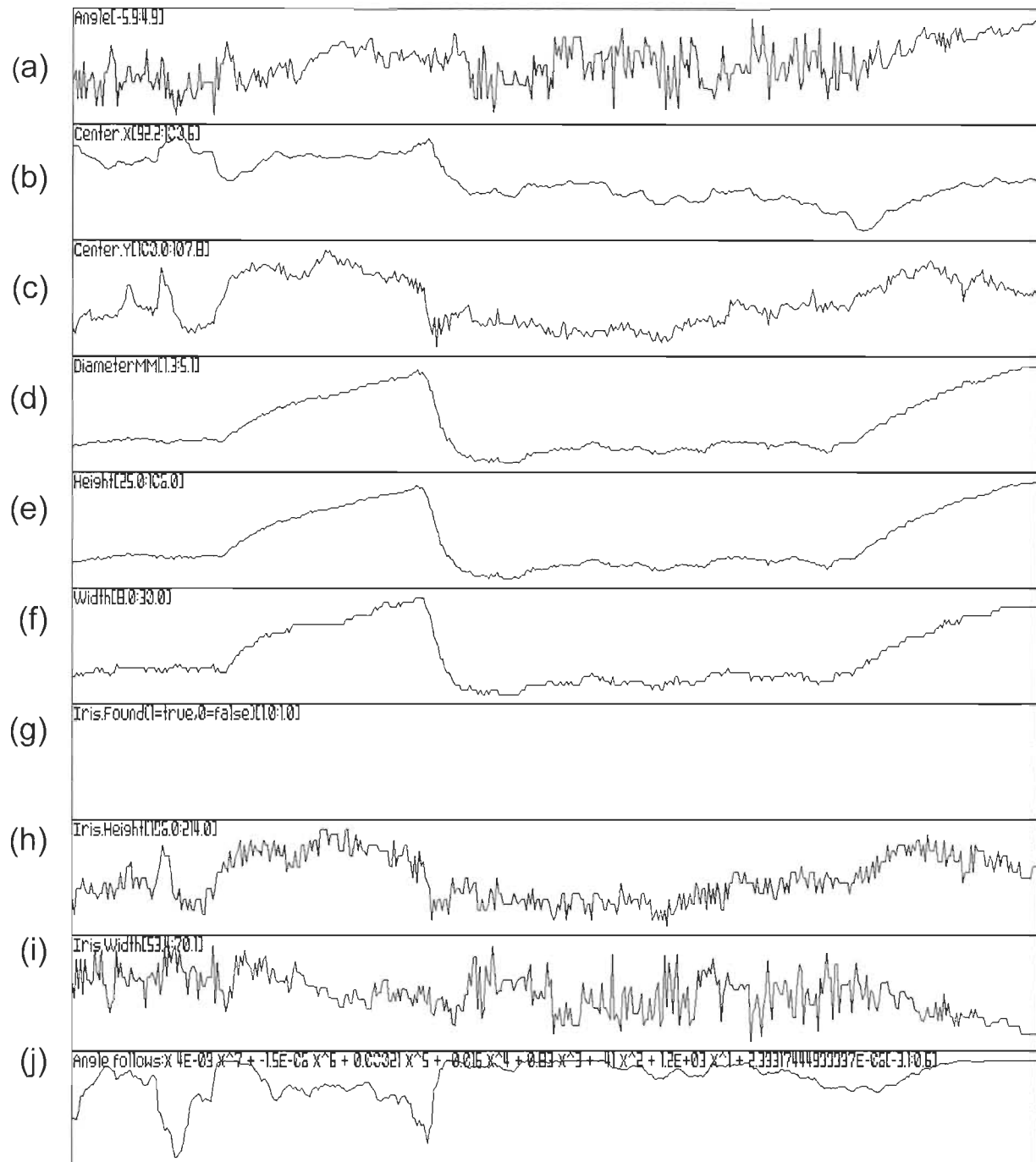


FIGURE 38 RÉSULTATS D'UN TEST DE DILATATION DE L'ŒIL GAUCHE

(a) Angle de visée (b) Position x de la pupille (c) Position y de la pupille (d) Diamètre de la pupille (e) Hauteur en pixel de la pupille (f) Largeur en pixel de la pupille (g) Si l'iris a été détecté dans cette image (h) Hauteur de l'iris détecté (i) Largeur de l'iris (j) Angle de visée en utilisant une polynomiale pour l'adoucir.

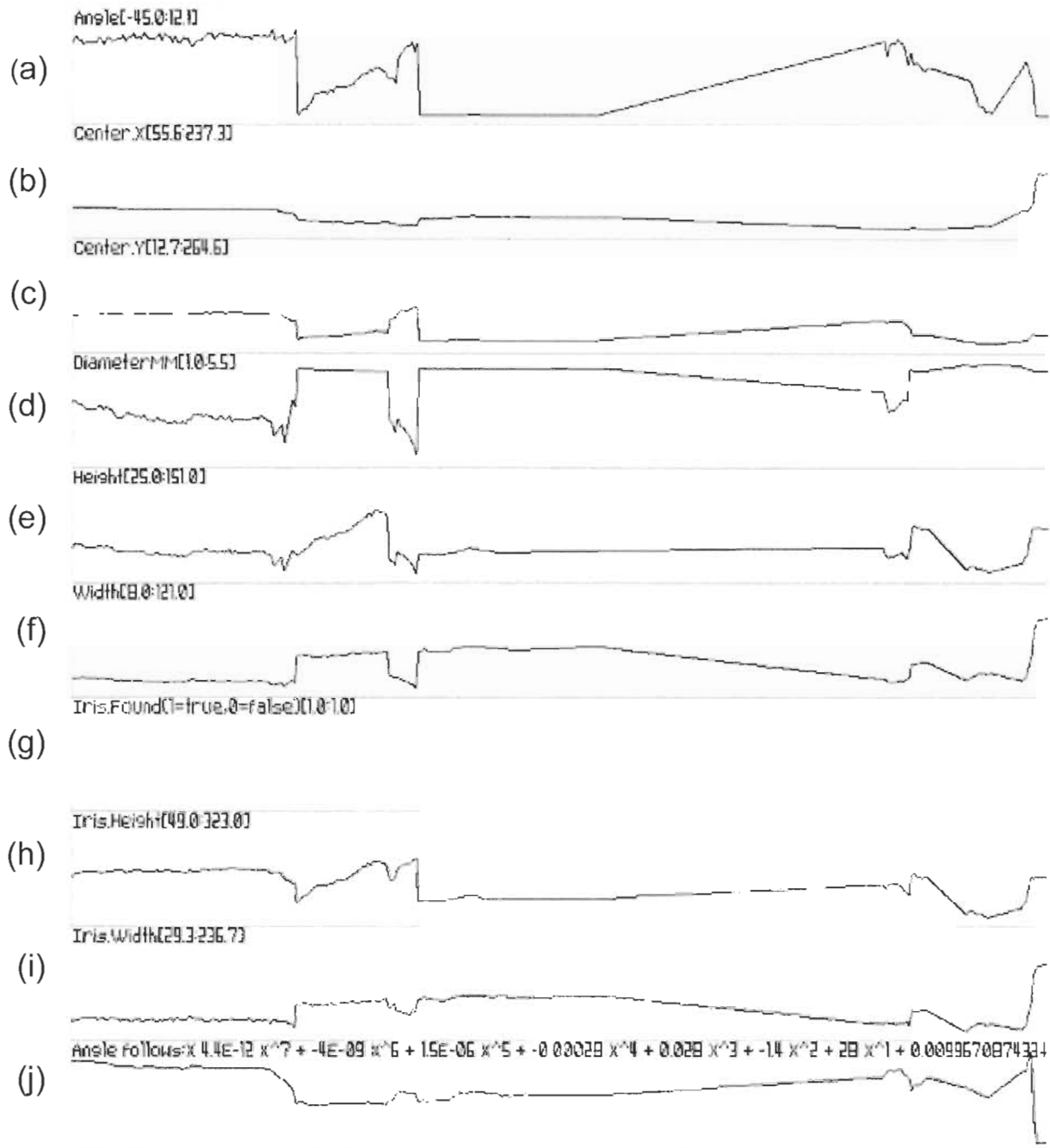


FIGURE 39 RÉSULTATS D'UN TEST DE NYSTAGMUS DE L'ŒIL GAUCHE

(a) Angle de visée (b) Position x de la pupille (c) Position y de la pupille (d) Diamètre de la pupille (e) Hauteur en pixel de la pupille (f) Largeur en pixel de la pupille (g) Si l'iris a été détecté dans cette image (h) Hauteur de l'iris détecté (i) Largeur de l'iris (j) Angle de visée en utilisant une polynômiale pour l'adoucir.

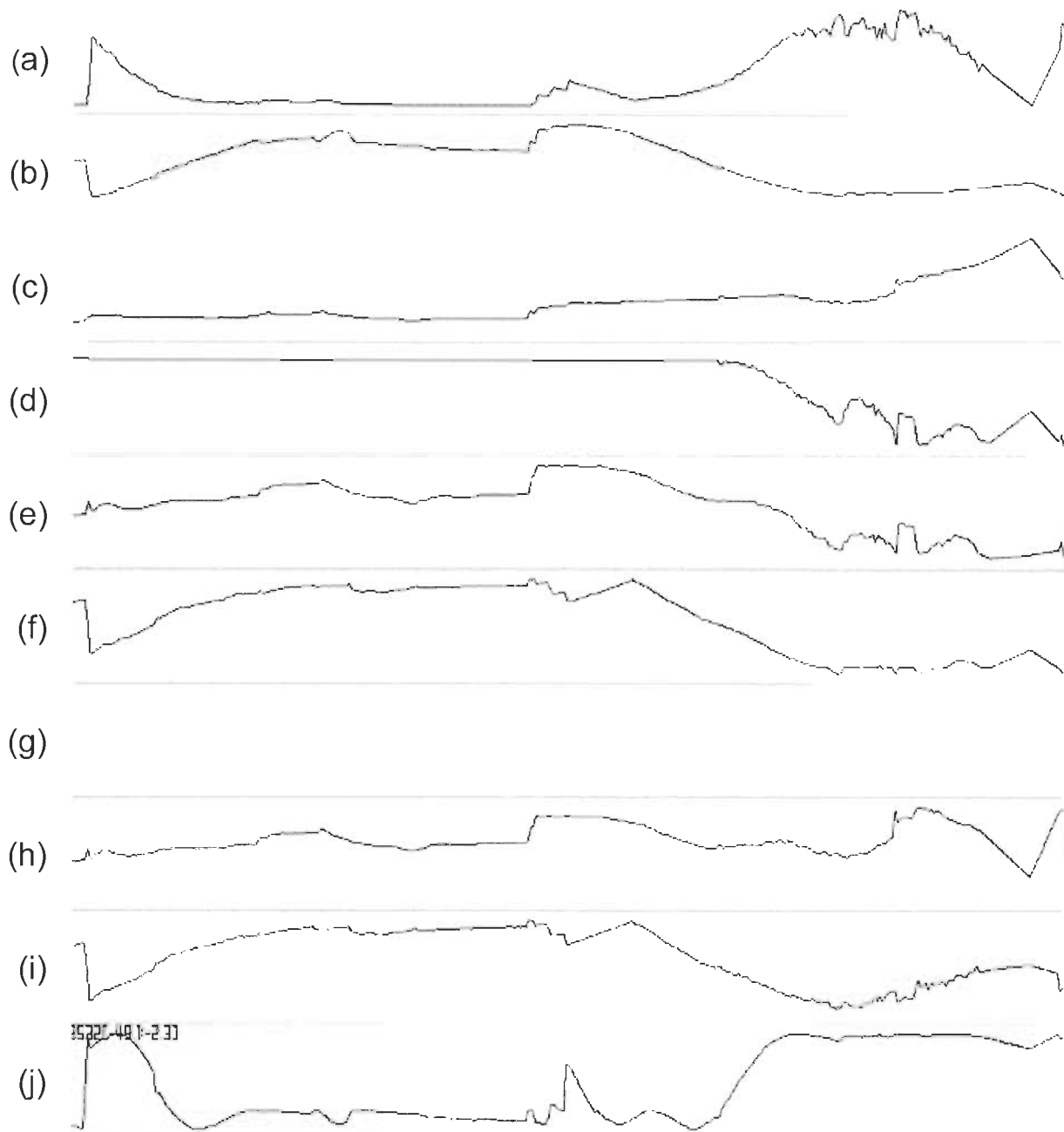


FIGURE 40 RÉSULTATS D'UN TEST DE NYSTAGMUS DE L'ŒIL GAUCHE(SUITE)

(a) Angle de visée (b) Position x de la pupille (c) Position y de la pupille (d) Diamètre de la pupille (e) Hauteur en pixel de la pupille (f) Largeur en pixel de la pupille (g) Si l'iris a été détecté dans cette image (h) Hauteur de l'iris détecté (i) Largeur de l'iris (j) Angle de visée en utilisant une polynomiale pour l'adoucir.

Pour le test de nystagmus horizontal gauche, il est possible d'extraire certaines informations pertinentes (voir Figure 39 et Figure 40), la valeur de centre en X (Figure 39 (b)) est faible sur la moitié du test, car l'œil regarde dans la direction opposée ce qui rend minimales les déplacements dus à l'angle de la caméra. Lors du suivi horizontal du stimulus visuel, l'œil regarde progressivement dans la direction de la caméra (Figure 40 (b)), une augmentation progressive de la valeur du centre en X apparaît représentant ainsi le déplacement horizontal effectué lors du suivi de l'objet 3D généré dans la scène 3D. Un phénomène imprévu se produit à la fin du test pour le centre en Y (Figure 40 (c)), car l'œil devrait suivre une ligne horizontale et ne pas grimper comme observé à la fin. Ce phénomène peut être attribué à une personne ne suivant pas parfaitement les consignes ou ayant eu un spasme oculaire. Il n'y a pas eu de dilatation de l'œil dans ce test puisque l'illumination était constante (Figure 40 (d)).

Pour faire suite à l'analyse de ces résultats, certaines décisions ont été prises avant d'en arriver à choisir uniquement 5 caractéristiques pour faire l'apprentissage automatique de la détection des facultés affaiblies, qui sont les 5 caractéristiques les plus proches des signaux bruts et les moins bruitées. Le choix a donc été porté sur le centre en X, en Y, le diamètre de la pupille en millimètres (celui-ci étant calculé en se référant sur la taille de l'iris et de la pupille), la largeur et la hauteur de la pupille en pixel. Le choix des caractéristiques les moins bruitées possible permet à l'algorithme de se concentrer sur les détails plutôt que sur le bruit et en sortir une meilleure détection. Une transformation des données n'appliquerait qu'une translation et un facteur d'échelle à la donnée, la translation est invariante dans l'algorithme et l'algorithme apprendrait seulement une courbe à l'échelle afin d'en sortir les mêmes résultats.

4.2 Résultats de l'apprentissage

Dans le Tableau 7, les colonnes CA, CB et CC représentent les choix qui ont dû être pris pour arriver à en faire l'apprentissage. Pour faire un rappel de ces choix; dans CA il était possible de choisir une séparation :

1. Par la dérivée numérique (changement de pente)
2. Par une durée fixe de 2 images
3. Par une durée fixe de 10 images
4. Par une durée fixe de 30 images

Dans CB le choix était un degré de liberté en 1 et 1000 par puissance de 10 (Distance maximale entre les mots).

1. 1
2. 10
3. 100
4. 1000

Pour le choix CC, il était possible de faire :

1. Une moyenne des données
2. Un minimum des données
3. Un maximum des données
4. Une multiplication des données

Chacune des autres colonnes (Tableau 1Tableau 7) représente un test, en d'autres mots, l'identification d'une catégorie de substances prise à une heure précise. Le titre de cette colonne est séparé en deux sections, la première est la catégorie de substance et la deuxième est la date et l'heure à laquelle ce test a été effectué (Catégorie : Date Heure).

Les cases du Tableau 7 représentent une valeur sur trois (3), alors si les textes extraits du Nystagmus, de la dilatation et de la convergence identifient correctement la catégorie, le score sera de trois (3). Si deux tests sur trois (2/3) donnent la bonne catégorie, il peut être conclu que la substance est réellement celle identifiée. En revanche si les trois tests représentent la même catégorie alors nos résultats seront sans aucun doute concluants.

C A	C B	C C	Alcool:201 9-03- 20_16h50	Alcool:2019 -03- 20_16h52	Alcool:2019 -03- 20_17h10	Alcool:201 9-03- 20_17h14	Normal:20 19-03- 20_16h32	Normal:2019- 04-01_14h24	Normal 19-01- 22_15h
4	1	4	2	1	1	0	3	2	3
4	1	1	2	1	1	0	3	1	3
4	1	3	2	1	1	0	3	1	3
4	1	2	0	0	0	0	3	3	3
4	2	2	0	0	0	0	3	3	3
4	3	2	0	0	0	0	3	3	3
2	1	2	0	0	0	0	3	3	3
2	2	2	0	0	0	0	3	3	3
2	3	2	0	0	0	0	3	3	3

...

Normal:20 19-02- 27_13h17	Normal:20 19-02- 27_13h32	Normal:201 9-03- 18_13h43	Normal:2 019-03- 18_13h4 7	Vidéo identifié (%)	Total	Vidéo identif ié %	2/3	3/3	Sujet identifié at 2/3 (%)
3	3	3	3	27	36	75	9	7	75
3	3	3	3	26	36	72	8	7	67
3	3	2	3	25	36	69	8	6	67
3	3	3	2	23	36	64	8	7	67
3	3	3	2	23	36	64	8	7	67
3	3	3	2	23	36	64	8	7	67
3	3	3	1	22	36	61	7	7	58
3	3	3	1	22	36	61	7	7	58
3	3	3	1	22	36	61	7	7	58

...

TABLEAU 7 RÉSULTAT D'IDENTIFICATION

CA	CB	CC	CD	Alcool:2019-03-20_16h50	Alcool:2019-03-20_16h52	Alcool:2019-03-20_17h10	Alcool:2019-03-20_17h14	Normal:2019-03-20_16h32	Normal:2019-04-01_14h24
4	1	1	1	A	A	A	N	N	A
4	1	1	4	A	A	A	N	N	A
4	1	1	9	A	A	A	N	N	A
4	1	1	12	A	A	A	N	N	A
4	1	3	1	A	A	A	N	N	A
4	1	3	3	A	A	A	N	N	A
4	1	3	4	A	A	A	N	N	A
4	1	3	9	A	A	A	N	N	A
4	1	3	11	A	A	A	N	N	A
4	1	3	12	A	A	A	N	N	A
4	1	4	1	A	A	N	N	N	N
4	1	4	3	A	A	N	N	N	N
4	1	4	4	A	A	A	N	N	A
4	1	4	9	A	A	N	N	N	N
4	1	4	11	A	A	N	N	N	N
4	1	4	12	A	A	A	N	N	A

...

SUITE PAGE SUIVANTE

CA	CB	CC	CD	Normal:2019-02-13_09h15	Normal:2019-02-27_13h17	Normal:2019-02-27_13h32	Normal:2019-03-18_13h43	Normal:2019-03-18_13h47
4	1	1	1	N	N	N	N	N
4	1	1	4	N	N	N	N	N
4	1	1	9	N	N	N	N	N
4	1	1	12	N	N	N	N	N
4	1	3	1	N	N	N	N	N
4	1	3	3	N	N	N	N	N
4	1	3	4	N	N	N	N	N
4	1	3	9	N	N	N	N	N
4	1	3	11	N	N	N	N	N
4	1	3	12	N	N	N	N	N
4	1	4	1	N	N	N	N	N
4	1	4	3	N	N	N	N	N
4	1	4	4	N	N	N	N	N
4	1	4	9	N	N	N	N	N
4	1	4	11	N	N	N	N	N
4	1	4	12	N	N	N	N	N
CA	CB	CC	CD	...				

TABLEAU 8 VALIDATIONS SELON LES 4 MODES (PARTIEL 83% ET PLUS)

	NORMAL	ALCOOL
NORMAL	116	12
ALCOOL	20	44

TABLEAU 9 VALIDATION CROISÉE

VALIDATION CROISÉE POUR LES MODES VU AU TABLEAU 8 (83% ET PLUS). FAUX POSITIFS VISIBLES EN NOIR ET FAUX NÉGATIFS EN GRIS.

	NORMAL	ALCOOL	NON-CLASSÉ
NORMAL	7	0	1
ALCOOL	1	3	0

TABLEAU 10 MATRICE DE CONFUSION

MATRICE DE CONFUSION POUR LES MODES CA=4, CB=1, CC=1, CD=9 ET CA=4, CB=1, CC=4, CD=11. FAUX POSITIFS VISIBLES EN NOIR ET FAUX NÉGATIFS EN GRIS. VALEURS RÉELLES À GAUCHE, VALEURS IDENTIFIÉE EN HAUT. EN CAS D'ACCORDANCE LA VALEUR EST CONSERVÉE SINON ILS SONT NON-CLASSÉS.

Dans le Tableau 7 présentant les résultats d'identification, il est possible de voir les modes CA, CB, et CC donnant les meilleurs résultats en prenant les tests séparément. Il est possible de voir que les tests normaux (sans facultés affaiblies) identifient souvent correctement les vidéos dans la bonne catégorie, soit normale, ou avec facultés affaiblies. Cependant, les tests après la prise d'alcool sont rarement dans la bonne catégorie. En ce qui a trait aux dernières colonnes du Tableau 7, ces colonnes avec les valeurs deux tiers (2/3) et trois tiers (3/3) représentent la possibilité de conclure sur la substance. Si au moins 2 ou 3 des trois tests (HGN, dilatation/contraction ou convergence) sont dans la bonne catégorie alors il est possible d'identifier cette catégorie. Un résultat de 75% n'est pas concluant étant donné que les résultats sont classés dans la catégorie normale (sans influence) trop souvent (voir colonnes alcool proches de 0/3).

Il est cependant possible de s'apercevoir que CA=4 revient très souvent et CB=1 aussi afin de donner les meilleurs résultats du système. Il faudrait possiblement voir si une division plus longue des séquences fixes pouvait donner un meilleur résultat. Rappelons que CA=4

représente une division des séquences en segments de 1 seconde et que le résultat est nettement supérieur aux tests sur un tiers de seconde ou d'un quinzième de seconde.

Dans le Tableau 8 décrivant la validation selon les 4 choix (CA, CB, CC, CD), il est possible d'apercevoir sur la gauche l'ensemble des possibilités de mode de l'application CA, CC et CD. Ensuite chaque colonne représente un individu, soit l'ensemble des tests du nystagmus, de dilatation et de convergence. Les cases représentent la catégorie dans laquelle ce mode a classé cet individu. La dernière colonne donne le nombre d'individus dans la bonne catégorie pour ce mode. Nous cherchons plus de 67% (2/3) d'identification, car cela est la proportion pour les tests normaux. Un taux inférieur nous indique uniquement que l'application est en mesure de classer l'ensemble des individus dans la catégorie contenant le plus de résultats.

Il a été possible de remarquer, grâce à ce tableau (Tableau 8), que 16 modes de calcul différents énumérés par les quatre premières colonnes sur les 768 modes (Combinaison des 4 possibilités CA, des 4 possibilités CB, des 4 possibilités CC et des 12 possibilités CD) donnent des résultats de 10/12 (Tableau 8 dernière colonne) ou 83%. Ce nombre est aussi corroboré par ce tableau montrant que dans certains cas le système répond trop souvent que les facultés sont normales uniquement deux fois sur les douze (2/12) cas où il identifie un individu en pleine faculté comme individu à facultés affaiblies par l'alcool. J'opterais alors pour le mode CA=4, CB=1, CC=1, CD=9 (Tableau 8). Ce mode répond plus souvent facultés affaiblies par l'alcool qu'il ne le devrait, pourrait permettre d'identifier les personnes n'ayant pas les facultés affaiblies dans un premier lieu. Il serait ensuite judicieux d'utiliser un mode classifiant plus de facultés normales tel que CA=4, CB=1, CC=4, CD=11 (Tableau 8). Celui-ci permettrait de bien identifier les individus qui sont vraiment en situation de facultés affaiblies par l'alcool. Pour ceux dont le résultat est différent dans les deux modes, il faudrait passer un test régulier du DRE, car le système ne serait pas en mesure de bien identifier la condition de facultés affaiblies. Ces cas seraient parmi les faux positifs et faux négatifs identifiés au Tableau 9. Le système est en mesure de reconnaître les faux positifs et faux négatifs et d'identifier qu'il n'est pas certains du classement de ceux-ci, éliminant ainsi partiellement les mauvaises détections (Tableau 10).

Pour donner suite à l'analyse du tableau complet contenant les 768 modes de calcul possibles, des tendances ont été remarquées. Cette première tendance se situe par rapport au mode CA (voir Figure 40). Le mode d'identification par dérivée, classe les individus dans la bonne catégorie en moyenne pour 33% de ceux-ci. Quant au mode par longueur constante, on parle respectivement de 44% pour une longueur d'un quinzième ($1/15$) de seconde, 48% pour un tiers ($1/3$) de seconde et 56% pour une coupure d'une seconde. Cela peut s'expliquer par le fait que plus la chaîne est longue, moins elle est portée au bruit pouvant identifier les sous-séquences comme différentes. Bien que 56% de bonne classification est assez faible, ce n'est qu'une moyenne. De plus, la tendance se maintient aussi pour les maximums. Un CA divisé par une très grande longueur fixe démontre de meilleurs résultats dû à une sensibilité inférieure au bruit.

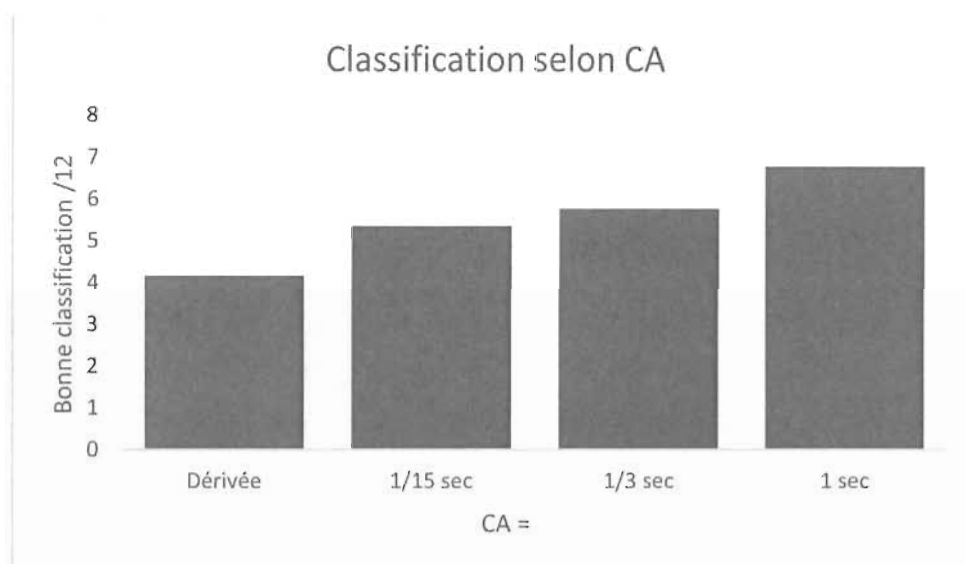


FIGURE 41 CLASSIFICATION SELON CA

En ce qui a trait au degré de liberté pour que deux mots soient similaires (CB), la tendance est contraire (Figure 42), soit que plus le degré est petit, meilleur est la classification. Rappelons que ce degré est exprimé par les puissances de dix. Cela s'explique par le fait de la grande différence entre 1, 10, 100 et 1000 comme degré de liberté, un CA offrant plus de liberté ne doit pas être jumelé avec un CB en offrant plus, car une trop grande liberté exprime plutôt la faible différence entre les différentes séquences et une classification de moins grande qualité.

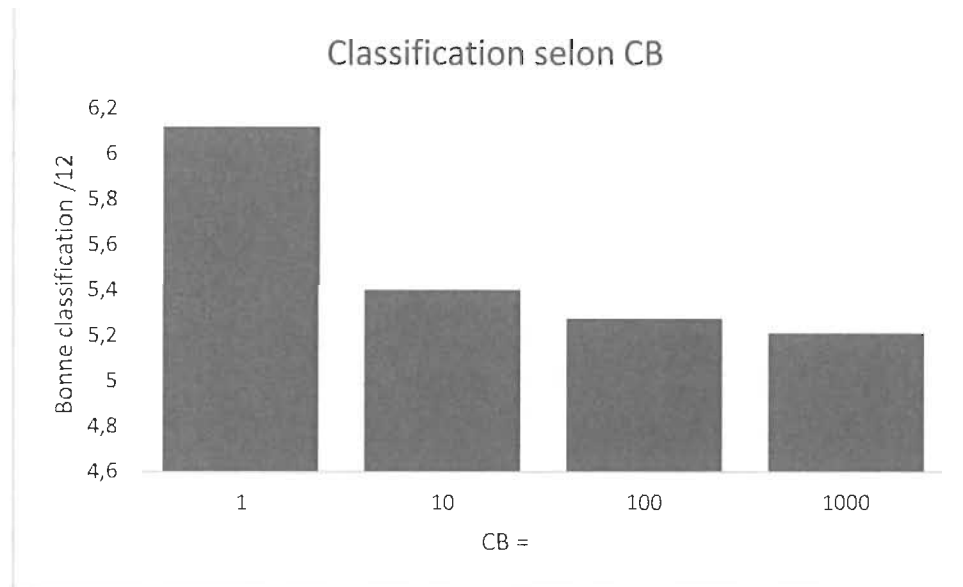


FIGURE 42 CLASSIFICATION SELON CB

Pour le choix d'unification des caractéristiques (CC), la tendance n'est pas linéaire dû au fait que les choix n'offrent pas un ordre quantitatif, mais qualitatif (Figure 43). Les meilleurs résultats s'obtiennent en appliquant le choix minimum. C'est-à-dire que parmi les données de similitudes entre les différents arbres, seule la valeur la moins ressemblante est prise en compte. Il est facile d'expliquer le bon résultat de ce mode en prenant en compte qu'un vidéo non ressemblant définit une différence, même si uniquement un des paramètres de ce vidéo change. Ce mode donne une classification de 55%, ce qui à première vue est extrêmement faible, cependant en prenant le meilleur des quatre choix, il sera possible d'extraire des taux de 83% comme vu précédemment. En ce qui a trait plutôt aux autres choix CC, la moyenne arithmétique obtenue donne un résultat de 45%, le maximum de ressemblance (distance entre les arbres) obtenu est de 46% et la multiplication des résultats donne 41%.

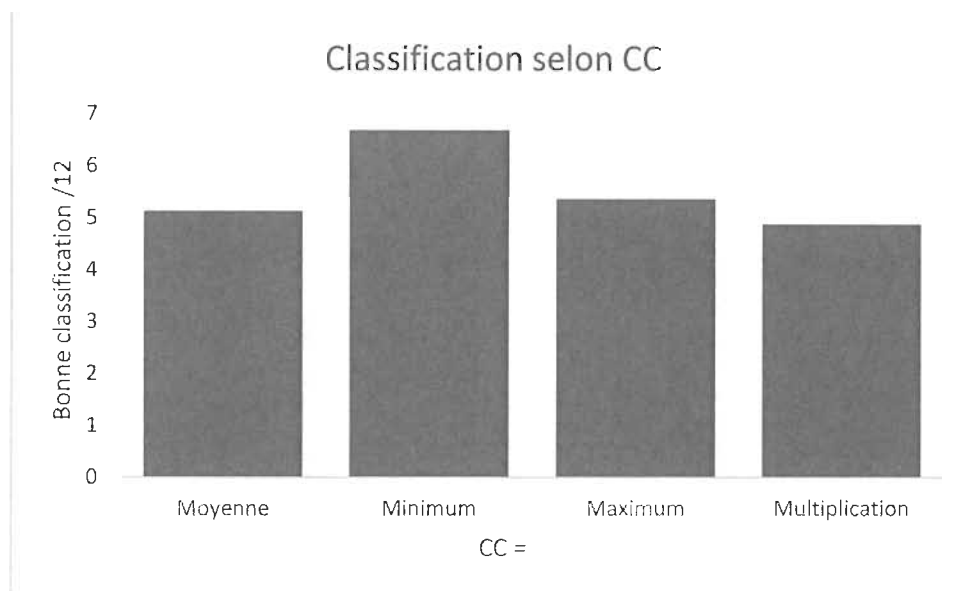


FIGURE 43 CLASSIFICATION SELON CC

Finalement pour la méthode jumelant les 3 tests (CD) (Figure 44), soit ceux du Nystagmus, de la dilatation et de la convergence, la tendance est plus difficile à extraire (Figure 44). Rappelons qu'il y a ici douze (12) choix possibles, ou plutôt quatre (4) avec trois (3) options possibles pour chacun de ces choix. Les choix de base sont encore la moyenne, le minimum, le maximum et la multiplication. Cependant dans ce cas spécial, il est aussi possible de conclure si certains nombres de tests donnent le même résultat. Dans ce cas les possibilités 5 à 8 représentent la même chose que les cas 1 à 4 en considérant que si deux des trois (2/3) tests donnent le même résultat il n'y a plus de doute sur le diagnostic des facultés d'un individu. La même chose s'applique aux cas 9 à 12, ceux-ci éliminent tout doute uniquement si les trois tests identifient la même chose. Il est cependant possible d'identifier que les cas 1 à 4 et 9 à 12 ont une moyenne similaire proche de 50%. Les cas 5 à 8 ont eu 44% de bonne classification. Cela peut s'expliquer par le fait qu'il est très rare que les trois tests identifient la même catégorie ce qui n'affecte pas vraiment les cas 9 à 12 qui se comportent alors identiquement à ceux 1 à 4. Dans les cas 5 à 8 cependant, il est possible de voir qu'il n'y a aucun changement entre les moyennes de ces quatre modes. Ce qui s'explique par le fait qu'il y a beaucoup de cas qui donnent deux résultats identiques.

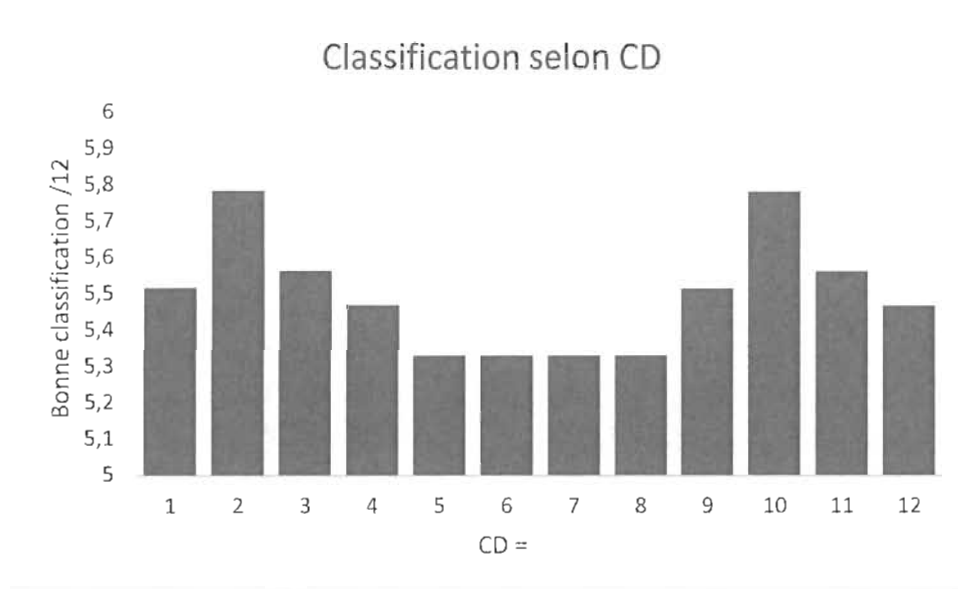


FIGURE 44 CLASSIFICATION SELON CD

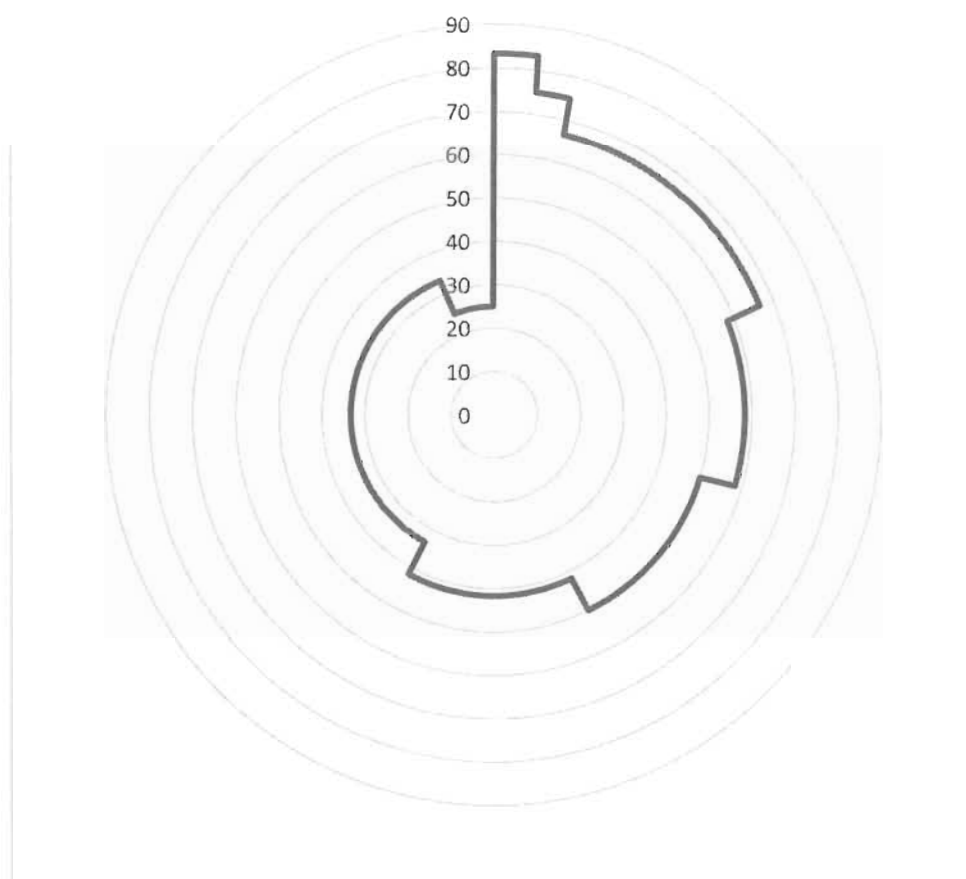


FIGURE 45 CLASSIFICATION DU NOMBRE DE MODES PAR LA QUALITÉ DE LA CLASSIFICATION EN %

Dans le graphique en radar (Figure 45), il est possible de s'apercevoir que les meilleures classifications, identifiées ici par les plus éloignés du centre, se concentrent dans très peu de mode. En regardant le tableau (Tableau 8) il est possible de voir que ces modes commencent par (4; 1; ...) soit la division par seconde avec un degré de liberté de seulement un. En regardant le tableau complet, il est alors possible d'identifier le mode (4; 3; 3; 1) comme donnant les pires résultats de classification avoisinant 25% de tous les modes. Il est aussi possible de remarquer la proportion de mode donnant des résultats acceptables est assez faible.

4.3 Discussion des résultats

La validité de mes résultats peut être mise en doute, car les tests ne portaient que sur une seule personne dans 12 situations différentes sur deux états. Cette situation ayant très peu de données ne me permettait pas de rejeter certains vidéos qui n'étaient pas aussi optimaux que prévu au niveau de leur qualité d'image (Voir Figure 46). Certains vidéos du casque étaient très difficiles à bien identifier par l'ordinateur et ont dû être analysés par un humain pour pouvoir tester l'algorithme de classification. Cette analyse manuelle a généré des imprécisions sur la localisation précise de l'œil dans les vidéos.

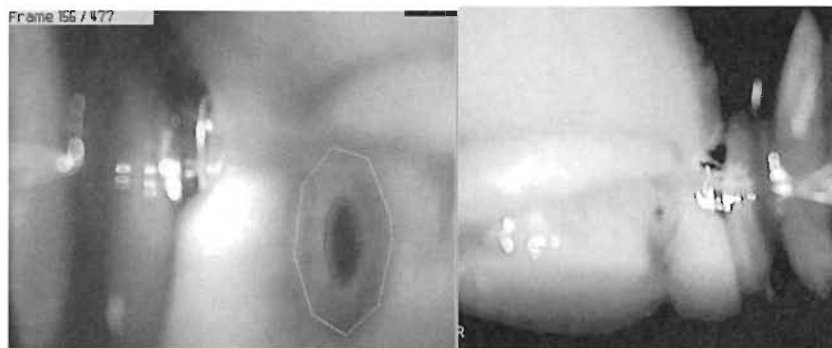


FIGURE 46 IMAGES D'UN ŒIL FLOU OU COUVERT

Dans tous les cas, une seule personne dans la banque d'apprentissage et seulement douze individus ne sont pas suffisants pour obtenir un résultat supérieur à 83% de bonne classification. Une taille de population dans la centaine pourrait certainement permettre un résultat de meilleure qualité.

Une amélioration possible de la détection serait d'utiliser le même principe appliqué par Cavnar [18], soit de retirer les sections communes entre les apprentissages. Dans notre cas ceci représente autant le temps d'attente juste avant le début du vidéo, par exemple, pour le test du nystagmus, du déplacement horizontal du stimulus visuel dans la scène 3D projeté dans le casque RV. Dans le cas où une certaine substance n'affecte pas un test donné, il serait judicieux de ne pas inclure la même courbe que si la personne n'avait pas consommé de substance, car

avoir deux catégories très similaires, ou dans ce cas identique ne peut que diminuer la performance du système.

4.4 Avancées possibles

Étant donné les résultats obtenus avec l'application de nos tests, il est possible de continuer le projet afin d'en améliorer la portabilité. Une des améliorations possibles serait d'utiliser une nouvelle version du casque de réalité virtuelle. Tout dernièrement, il a été annoncé qu'Oculus produirait une nouvelle version du casque nommée « Oculus Rift S » [29] qui serait plus abordable et moins encombrante que celle utilisée pour obtenir les résultats présentés dans ce mémoire. Ce nouveau casque semble prometteur sur beaucoup de points puisqu'il ne comprend qu'un seul port USB au lieu de deux et aurait un confort plus important. Au lieu d'utiliser un socle déposé sur une table afin de se positionner, une gamme de cinq caméras est utilisée pour faire la modélisation 3D de la pièce ambiante. Bien qu'il ne comporte pas de caméras intérieures, il semble qu'il puisse être possible d'insérer des caméras dans la même position que le modèle utilisé actuellement. Mais n'ayant pas la distance entre les yeux ajustable, il se peut qu'il utilise plutôt une surface solide plutôt que du tissu intérieur comme actuellement. C'est ce tissu qui a été retiré pour pouvoir insérer nos caméras. Sans ce tissu, il faudrait plutôt coller les caméras avec des socles fixes puisqu'aucun orifice ne pourrait supporter les caméras en place tel que présentement.

L'utilisation d'un port Display Port au lieu d'un port HDMI pourrait nuire cependant au projet, la majorité des ordinateurs « VR Ready » ont présentement uniquement des ports HDMI, car ces casques utilisent ce genre de connecteur. Il faudrait alors plutôt trouver un ordinateur avec une carte vidéo compatible et le port adéquat pour permettre son intégration dans une voiture de police. Ce nouveau casque est actuellement annoncé à la moitié du prix du dispositif original ce qui permettrait de rendre le tout plus abordable ne requérant pas nécessairement un poste de travail fixe. Uniquement, un ordinateur portable, le casque RV et une alimentation pour nos caméras NIR seraient alors nécessaires pour faire fonctionner le dispositif. Dans le passé, il fallait avoir un ordinateur, le capteur de bureau (pour positionner le casque), le casque de réalité virtuelle en plus de l'alimentation des caméras. Étant donné que les caméras requièrent des ports USB sur des contrôleurs séparés, un ordinateur portable ne pouvait généralement pas suffire,

puisque le casque nécessitait lui-même deux ports USB. Avec la nouvelle version, le casque à une seule connexion USB permettrait alors de connecter les deux caméras NIR, une de chaque côté de l'ordinateur portable pour permettre les mouvements des deux yeux en même temps plutôt que de passer le test en deux phases, une par œil.

4.5 Conclusion

Les résultats présentés précédemment montrent des perspectives très encourageantes pour le prototype développé dans cette recherche. Bien que n'étant pas parfaits pour l'instant, ils sont proches de l'être et la tendance pointe vers le fait qu'une augmentation des données d'apprentissage pourrait permettre de mieux apprendre concernant les zones où les différents modes choisis ne classent pas dans la même catégorie. La tendance de l'algorithme à savoir quels résultats ne peuvent être classés avec certitude nous démontre qu'il est conscient des données et de cette frontière peu définie. De plus amples tests pourraient être effectués pour lui apprendre cette frontière, mais la preuve de concept a été faite qu'il est possible d'utiliser cet algorithme pour la détection des facultés affaiblies. Il suffirait d'ajouter un apprentissage pour chaque catégorie de drogue pour que l'algorithme en fasse la détection.

CHAPITRE 5

CONCLUSION

L'objectif principal de cette recherche était de pouvoir effectuer les tests du nystagmus, de dilatation/contraction et de convergence dans un environnement contrôlé pour éviter les erreurs possibles causées par un changement d'environnement. Rappelons que le test de dilatation/contraction nécessite un environnement des plus contrôlés sinon l'œil n'aura pas une taille connue et stable au repos. Le test de nystagmus pourrait quant à lui être altéré si un objet autre que celui que nous devons suivre est présent pouvant causer des distractions, de même que pour la convergence qui nécessite aussi qu'aucune distraction ne soit présente. Le casque de capture installé à l'intérieur d'un Oculus Rift a été une réussite mitigée. L'observation des vidéos capturées par les caméras NIR installées dans le casque RV sont floues, mais assez visibles pour permettre d'identifier la pupille et l'iris. Une segmentation en utilisant K-Means est suffisante pour rendre l'image utilisable. Cependant, un algorithme plus robuste serait requis pour mieux segmenter et donc en améliorer l'identification, il y a souvent de faux positifs occasionnés par l'ombrage du nez dû au positionnement de la caméra et au fait que l'on perd l'œil quand la direction de regard est dans la direction opposée. Il était alors prévu que le manque d'information était une information en soi, mais l'occurrence de faux positifs la rendent inutilisable. Cependant, ce nouveau prototype, comme preuve de concept, a permis de produire des analyses qui ont été un succès pour certains vidéos alors il ne faudrait surement que de petits ajustements pour en améliorer les performances.

Le principal aspect novateur du dispositif physique est l'utilisation d'un casque de réalité virtuelle, principalement conçu pour les jeux en immersion afin d'aider les agents de la paix dans la détection des cas de facultés affaiblies et leur permettre de rendre les routes plus sécuritaires.

Un autre objectif était que le logiciel utilise un algorithme en arbre basé sur les N-Gram généralement utilisé pour les textes afin de classer les individus dans la bonne catégorie d'état de leurs facultés. Avec beaucoup de mode testé, il a été possible d'en identifier seize pouvant donner un résultat utilisable pour la détection. Je crois qu'une détection de 83% donne une preuve

de concept réussie. Cet algorithme démontre une tendance à pouvoir détecter différentes catégories de facultés d'un individu malgré la faible quantité de données lui ayant été présentées, offrant ainsi la possibilité d'avoir une banque d'apprentissage de taille raisonnable. La performance de l'algorithme est aussi très rapide ce qui permettrait de faire les tests sans nécessiter un temps d'attente considérable ou un appareil performant. Il manque effectivement beaucoup de travail sur les détails tel que l'identification automatique des modes à utiliser pour la classification et sur la quantité de données, mais un taux de réussite aussi élevé démontre que le système a bien appris. Le 17% restant pourrait être remis à des tests plus traditionnels, diminuant la demande pour ceux-ci. Ce 17% est identifié par un résultat différent lorsqu'on exécute plusieurs modes de calcul parmi les seize énoncés.

BIBLIOGRAPHIE

- [1] Société de l'assurance automobile du Québec, [En ligne]. Available: <http://www.saaq.gouv.qc.ca/prevention/alcool/chiffres.html>.
- [2] NHTSA, Drug evaluation and classification training program HS 172, R4/93.
- [3] J. Stallkamp, M. Schlipsing, J. Salmen et C. Igel, «Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition,» 16 Février 2012. [En ligne]. Available: <https://www.sciencedirect.com/science/article/pii/S0893608012000457>.
- [4] P. DROLET, DÉTECTION AUTOMATIQUE DU NYSTAGMUS CHEZ UN INDIVIDU SOUS INFLUENCE D'ALCOOL, TROIS-RIVIERES, QUEBEC: UQTR, 2004.
- [5] P.-O. HOUDE, DÉTECTION DE FACULTÉS AFFAIBLIES PAR L'ANALYSE DU COMPORTEMENT OCULAIRE SUR DES SÉQUENCES VIDÉO, TROIS-RIVIERES, QUEBEC: UQTR, 2008.
- [6] D. ADAM, AUTOMATISATION ET ANALYSE DES RÉSULTATS DE TESTS SYMPTOMATIQUES LIÉS AU COMPORTEMENT OCULAIRE POUR LA DÉTECTION DE FACULTÉS AFFAIBLIES PAR L'ALCOOL OU LES DROGUES, TROIS-RIVIERES, QUEBEC: UQTR, 2012.
- [7] D. LAPERRIÈRE, AUTOMATISATION DE TESTS SYMPTOMATIQUES LIÉS AU COMPORTEMENT OCULAIRE POUR LA DÉTECTION DE FACULTÉS AFFAIBLIES PAR L'ALCOOL OU LES DROGUES, TROIS-RIVIERES, QUEBEC: UQTR, 2009.
- [8] V. LECOURS, DÉTECTION AUTOMATIQUE DES FACULTÉS AFFAIBLIES SUR APPAREILS MOBILES, TROIS-RIVIERES, QUEBEC: UQTR, 2015.

- [9] Guide Vue, «L'anatomie de l'œil,» 26 Aout 2018. [En ligne]. Available: <https://www.guide-vue.fr/la-sante-de-vos-yeux/shemas-de-l-%C5%93il>.
- [10] SAAQ, «Chapitre C-24.2,» chez *Code de la sécurité routière*, Éditeur officiel du Québec.
- [11] American Addiction Centers, "Signs of Drug Use in the Eyes: Pupil Dilation and Redness," [Online]. Available: <https://americanaddictioncenters.org/health-complications-addiction/signs-drug-use-eyes>.
- [12] E. Schechtman et D. Shinar, «Modeling drug detection and diagnosis with the 'drug evaluation and classification program',» chez *Accident Analysis and Prevention* 37, Beer sheva, Israel, 2005.
- [13] EMGU.CV, «EMGU.CV,» 10 Mai 2008. [En ligne]. Available: http://www.emgu.com/wiki/index.php/Main_Page.
- [14] Moyskley Tech., «ImageProcessing,» 22 Novembre 2017. [En ligne]. Available: <https://github.com/moyskleytech/ImageProcessing>.
- [15] N. Dhanachandra, K. Manglem et Y. J. Chanu, «Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm,» chez *Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015)*, Manipur, India, 2015.
- [16] Oxford Dictionaries, «Definition of n-gram in English,» [En ligne]. Available: <https://en.oxforddictionaries.com/definition/n-gram>.
- [17] P. F. Brown, P. V. deSouza, R. L. Mercer, V. J. Della Pietra et J. C. Lai, «Class-Based n-gram Models of Natural Language(Revised),» *Computational Linguistics*, vol. 18, n° %14, pp. 467-479, 1992.
- [18] W. Cavnar, «N-Gram-Based Text Categorization,» *Proceedings of SDAIR*, 1994.

- [19] V. Kešelj, F. Peng†, N. Cercone† et C. Thomas, «N-GRAM-BASED AUTHOR PROFILES FOR AUTHORSHIP ATTRIBUTION,» *Pacific Association for Computational Linguistics*, 2003.
- [20] T. Niesler et P. Woodland, «A variable-length category-based n-gram language model,» *IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, 1996.
- [21] J. Furnkranz, «A Study Using -gram Features for Text Categorization,» Austrian Research Institute for Artificial Intelligence, Wien, Austria, 1998.
- [22] P. F. Brown, P. V. deSouza, R. L. Mercer, V. J. Della Pietra et J. C. Lai, «Class-based ngram models of natural language,» 1990.
- [23] A. Pauls et D. Klein, «Faster and Smaller N-Gram Language Models,» *Computer Science Division*.
- [24] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu et P. Kuksa, «Natural Language Processing (Almost) from Scratch,» *Journal of Machine Learning Research*, 2012.
- [25] G. C. Cawley, «Leave-One-Out Cross-Validation Based Model Selection Criteria for Weighted LS-SVMs,» chez *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, Vancouver, BC, Canada, 2006.
- [26] J. A. H. a. M. A. Wong, «Algorithm AS 136: A K-Means Clustering Algorithm,» *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 28, n° 11, pp. 100-108, 1979.
- [27] Silicon Labs, «SI8902B - Datasheet,» 2012.
- [28] «Easy Pulse Plugin: An Arduino pulse sensor,» [En ligne]. Available: <https://www.tindie.com/products/rajbex/easy-pulse-plugin-an-arduino-pulse-sensor/>.
- [29] Upload VR, «Oculus Rift S Official,» 2019. [En ligne]. Available: <https://uploadvr.com/oculus-rift-s-official/>.

- [30] Comité des analyses d'alcool, «Normes et procédures recommandées par la Société canadienne des sciences judiciaires,» *Canadian Society of Forensic Science*, vol. 36, n° 13, pp. 129-159, 2003.
- [31] NHTSA, DWI Detection and standardized field sobriety testing. HS 178, R6/92.
- [32] [En ligne]. Available: <https://www.adafruit.com/product/2699>.